Digital Signal Processing 49 (2016) 65–85

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Blind adaptive carrier phase recovery for QAM systems

Shafayat Abrar [a,*], Azzedine Zerguine [b], Asoke K. Nandi [c,1]

[a] COMSATS Institute of Information Technology, Islamabad 44000, Pakistan
[b] King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia
[c] Brunel University London, Uxbridge, Middlesex UB8 3PH, United Kingdom

ARTICLE INFO

Article history:
Available online 10 November 2015

Keywords:
Adaptive receiver design
Blind algorithm
Carrier phase recovery
Quadrature amplitude modulation
Quadrature phase shift keying
Mean square deviation

ABSTRACT

This paper provides a study of adaptive phase recovery in quadrature amplitude modulation (QAM) based communication systems. Here, we modify the traditional fourth-power phase recovery algorithm (FP-PRA) to propose three improved algorithms, and analyze their performances in the presence of additive white Gaussian noise, phase noise, frequency-offset, and inter-symbol interference. We demonstrate that it is possible to obtain the optimal values of step-sizes (or loop-gains) in closed-form in the presence of phase noise and/or frequency-offset. In particular, we discuss two methods to improve FP-PRA. The first method involves utilizing the idea of partitioning the QAM constellation into QPSK-like and not-QPSK-like annular regions. The phase synchronizer is allowed to update only when a derotated QAM symbol lies in QPSK-like region; otherwise, the update process is stopped. The second method exploits an evolving idea of QAM-to-QPSK transformation, and uses transformed symbols to estimate phase mismatch. We provide a new interpretation of this transformation method and relate it to the quadrant-wise centroid of the rotated constellation. Furthermore, we discuss the feasibility of this method for both square and cross-QAM, and, identify and verify numerically the existence of false-locks in the case of cross-QAM. To the best of our knowledge, the ideas of constellation partitioning and constellation transformation have not appeared earlier in the context of adaptive phase recovery. We discuss adaptive blind estimation of optimal step-sizes in the presence of phase noise and frequency-offset. Finally, we discuss the modification of the proposed stochastic gradient methods to transform them into batch processing algorithms so as to make them more suitable for higher data rate systems. Numerical experiments indicate that the proposed algorithms can outperform the traditional FP-PRA algorithm for a number of practical QAM sizes under different mismatch conditions, and that our analytical findings are in close agreement with the simulation results.

1. Introduction

The recovery of digital transmissions without resorting to training signals is an important problem due to the increasing interest in digital broadcasting and communications. In the literature, this is commonly referred to as *blind adaptive receiver* since it involves blind equalization, i.e., removing inter-symbol interference, as well as carrier phase tracking, time recovery, and gain control [1]. A quadrature amplitude modulation (QAM) based digital communication is a bandwidth efficient transmission technique. It makes use of multiple signal phase and amplitude levels to carry multiple bits per symbol. However, it requires accurate carrier phase and frequency estimation at the receiver to satisfy the need of successful signal demodulation and data detection. These parameters, for efficiency reasons, often have to be estimated with no knowledge of the transmitted preamble sequence [2].

This work aims to discuss adaptive recovery of carrier-phase in QAM-based digital communication systems. This is organized as follows: Section 2 presents the baseband model of the system, the basic properties of sensitivity (error) function for phase-offset detection, and a generic adaptive synchronizer for the correction. Section 3 revisits the well-known adaptive decision-directed and the fourth-power phase recovery algorithms, and also sheds some light on the stationary points. Section 4.1 presents the first proposal to partitioning the QAM constellation into QPSK-like and not-QPSK-like regions to enhance the phase recovery capability, and which allows the update only when the derotated symbols lie in QPSK-like regions (a *stop-and-go* policy). Section 4.2 presents the second proposal to transforming the QAM constellation into a QPSK constellation so that the phase update is made for every

* Corresponding author.
E-mail addresses: sabrar@comsats.edu.pk (S. Abrar), azzedine@kfupm.edu.sa (A. Zerguine), asoke.nandi@brunel.ac.uk (A.K. Nandi).
[1] Distinguished Visiting Professor, Tongji University, Shanghai, China.

http://dx.doi.org/10.1016/j.dsp.2015.10.009

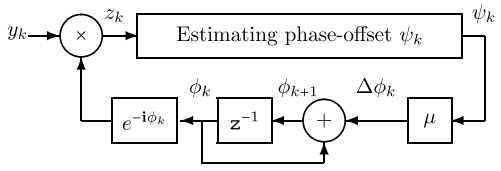

Fig. 1. An adaptive blind phase synchronizer. $\psi_k = \theta - \phi_k$ is the true error which has to be estimated from z_k. The μ is the loop-gain.

QAM alphabet by simply estimating the angular distance of the transformed symbol from the diagonal of the respective quadrant or fourth-power phase estimate (an *always-go* policy). This section also discusses the admissibility of the transformation method for phase recovery in square/cross-QAM systems. Section 5 evaluates the proposed algorithms for both steady-state acquisition and tracking abilities in the presence of phase noise, frequency-offset, and inter-symbol-interference. Section 6 discusses adaptive methods to obtain optimal values of loop-gain (step-size) for a system suffering with phase noise and frequency-offset. Section 7 discusses an adaptive but a batch processing algorithm for phase recovery. Simulation results and conclusions are presented in Sections 8 and 9, respectively.

2. Problem formulation

Consider a baseband QAM communication system in which x_k represents the transmitted data and y_k indicates the received signal. We assume that x_k is an independent identically distributed (i.i.d.) sequence, and its in-phase and quadrature components are of identical and symmetrical distributions. Furthermore, we assume initially that the system is already equalized and frequency synchronized, and the timing recovery has been established; so we focus solely on carrier phase recovery.[2]

Under the said assumptions, the baud-rate samples of the output of the matched filter (at the receiver end) are described by

$$y_k = \exp(\mathbf{i}\theta)x_k + w_k,\ k = 1, 2, 3, \cdots \tag{1}$$

where k is the discrete index denoting the time kT, $1/T$ is the signaling rate, θ denotes the unknown phase to be estimated and w_k is circularly symmetric i.i.d. zero-mean Gaussian random variable with independent real and imaginary parts [3]. The blind detection problem is to find an estimate of θ, without actually detecting the data x_k. Let ϕ_k be an estimate of θ available at the receiver, then an estimate of the original signal x_k is given by

$$z_k = \exp(-\mathbf{i}\phi_k)y_k. \tag{2}$$

As the actual input x_k is unavailable, different minimization criteria are usually explored. Such minimization cost functions do not directly involve the input signal x_k but still reflect the residual phase-offset in the synchronizer output. Define the mean cost function as

$$J(\phi) = \min_{\phi} \mathbb{E}\mathcal{J}(z_k), \tag{3}$$

where $\mathcal{J}(z_k)$ is a scalar function of z_k. Thus the design of the blind synchronizer translates into the selection of a suitable function $\mathcal{J}(z_k)$ such that the local minima of $J(\phi)$ corresponds to a significant or complete removal of the phase-offset from z_k. Refer to Fig. 1 for an architecture of a feedback (adaptive) phase-offset synchronizer.

[2] The adverse effects of residual inter-symbol interference (due to imperfect equalization), phase noise (due to the instability of oscillators) and frequency-offset (due to mismatch in local oscillators) on the performance of phase-recovery loop are addressed as special cases in Section 5.

Defining $\theta = \phi + \psi$, where ψ is the estimation error (or residual phase offset), we obtain

$$z_k = y_k \exp(-\mathbf{i}\phi_k) = x_k \exp(\mathbf{i}\psi_k) + e_k, \tag{4}$$

where e_k has similar statistical properties as those of additive noise w_k; refer to (1). Let $\mathcal{S}(\psi)$ be the sensitivity function (also called phase detector characteristics [4] or loop *S*-curve [5]), defined as[3]

$$\mathcal{S}(\psi) = -\frac{\partial J(\psi)}{\partial \psi} = -\mathbb{E}_{x,e}\left[\frac{\partial \mathcal{J}(\psi)}{\partial \psi}\middle|\psi\right], \tag{5}$$

where $\mathbb{E}_{x,e}[\cdot]$ denotes the statistical average over the data and noise. Using (3)–(5), the stochastic-approximation gradient-descent (minimization) algorithm is easily derived as

$$\phi_{k+1} = \phi_k - \mu\frac{\partial \mathcal{J}(z_k)}{\partial \phi}\bigg|_{\phi=\phi_k} = \phi_k + \mu\mathcal{S}(\psi_k), \tag{6}$$

where $0 < \mu \ll 1$ is a constant usually termed as the step-size or the loop-gain. $\mathcal{S}(\psi)$ should satisfy the following properties [4]:

1. $\mathcal{S}(\psi) = \mathcal{S}\left(\psi + n\frac{\pi}{2}\right),\ n \in \mathbb{Z}$
2. $\mathcal{S}(-\psi) = -\mathcal{S}(\psi)$
3. $\psi = 0$ is the only stable solution of $\mathcal{S}(\psi) = 0$ over $\left[0, \frac{\pi}{2}\right]$.

Property 1 arises from the quadrant symmetry of QAM signal sets (ambiguity due to this symmetry can be resolved with differential encoding). Property 2 is required as a sufficient condition for having a stable equilibrium at the origin ($\psi = 0$). Property 3 is required to avoid false lock points (or spurious minima). A corollary of Property 3 is that $\mathcal{S}(\psi)$ should be unimodal in $-\pi/4 < \psi \le \pi/4$, which implies that it is only possible to recover θ in the range $|\theta| \le \pi/4$. In the light of these properties, an ideal sensitivity function $\mathcal{S}(\psi)$ or the corresponding ideal cost $J(\psi)$ should exhibit following properties:

4. $J(\psi)$ should have the following stationary points and no saddle points:

   Local minima at $\psi = n\dfrac{\pi}{2},\quad n = 0, \pm1, \pm2, \pm3, \cdots$

   Local maxima at $\psi = n'\dfrac{\pi}{4},\quad n' = \pm1, \pm3, \pm5, \cdots$

   Sensitivity is zero at these stationary points.
5. For all ψ, other than the stationary points, $\mathcal{S}(\psi)$ should be proportional to ψ, that is $\mathcal{S}(\psi) = c_x\psi$, where c_x is constant of proportionality which possibly depends on the statistics of the signal.

Property 4 is a generalization of Property 1 and 3 to incorporate local minima and maxima. Property 5 is based on experimental evidence that, irrespective of the target phase to acquire, the higher the value of sensitivity in proportion to the phase-offset, the higher is the rate of convergence with minimum steady-state error. Property 5 is based on the findings of reference [6]. Satisfying Properties 1–5, a possible candidate for an ideal sensitivity function could be the one as illustrated in Fig. 2.

However, an ideal sensitivity function is closely possible to attain only for QPSK signal where we have one alphabet on the

[3] We borrow from [6] that the sensitivity function of an adaptive element is defined to be the negative of derivative of the minimizing cost with respect to the parameter of interest. If $\mathcal{S}_\zeta = -\partial J(\zeta)/\partial\zeta$ is the sensitivity function for ζ then ζ can be updated as $\zeta_{k+1} = \zeta_k + \mu\mathcal{S}_\zeta$. Since ζ is updated proportionally to $\mu\mathcal{S}_\zeta$, where $\mathcal{S}_\zeta$ determines the speed at which ζ can change.

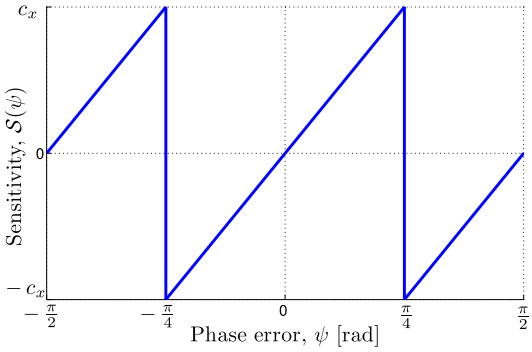

Fig. 2. An ideal sensitivity function for phase synchronization.

diagonal of each quadrant. In the presence of non-zero phase-offset, the angular deviation of the given alphabet away from the respective diagonal can yield an estimate of the unknown phase-offset; the resulting estimate would be proportional to $\sin(\psi_k)$ for $|\psi_k| \leq \pi/4$, refer to [7,8] for details.

3. Existing adaptive phase recovery algorithm

The earliest adaptive phase recovery algorithms were decision-directed in nature. Like Kobayashi [9], Simon and Smith [10] devised (independently) the following adaptive algorithm in the year 1971 and 1974, respectively:

$$\phi_{k+1} = \phi_k + \mu\, \Im[z_k \widehat{x}_k^*], \tag{7}$$

where $\widehat{x}_k$ is the decision made on z_k and $\Im[\cdot]$ is used to denote imaginary part of the enclosed complex entity. The resulting sensitivity function of (7) is linear only if $|\psi_k| < \theta_c$, where θ_c is the critical angle beyond which a QAM receiver mostly makes erroneous decisions. Owing to Mathis [11], the critical angle is obtained as:

$$\theta_c = \tan^{-1}\left(\frac{d}{(2L_M - 1)d}\right) = \tan^{-1}\left(\frac{1}{2L_M - 1}\right), \tag{8}$$

where L_M denotes the number of unique positive voltage levels of QAM, d and $(2L_M - 1)d$ are the smallest and highest amplitude in the constellation, respectively (usually, we consider $d = 1$). The L_M is obtained as $\sqrt{M}/2$ and $3\sqrt{M/32}$ for square- and cross-QAM, respectively. Beyond θ_c, the sensitivity function of (7) is found to exhibit deep false-locks. To reduce the depth of false-lock, Falconer suggested the following improved update [12] in the year 1976:

$$\phi_{k+1} = \phi_k + \mu\, \frac{\Im[z_k \widehat{x}_k^*]}{|z_k|\,|\widehat{x}_k|}. \tag{9}$$

However, both updates suffered with deep false-locks for cross-QAM signals. During 80's, Leclert and Vandamme [4], and Moridi et al. [13–15] came up with a plethora of decision-directed sensitivity functions for adaptive phase recovery, some of those were[4] $\Im[z_k \mathrm{csign}[\widehat{x}_k - z_k]^*]$, $\Im[\mathrm{csign}[z_k](\widehat{x}_k - z_k)^*]$, $\Im[\widehat{x}_k \mathrm{csign}[\widehat{x}_k - z_k]^*]$, $\Im[\mathrm{csign}[z_k]\mathrm{csign}[\widehat{x}_k - z_k]^*]$, $\Im[\widehat{x}_k(\widehat{x}_k - z_k)^*]$, $\Im[\mathrm{csign}[\widehat{x}_k](\widehat{x}_k - z_k)^*]$, and $\Im[\mathrm{csign}[\widehat{x}_k]\mathrm{csign}[\widehat{x}_k - z_k]^*]$. However, all of these sensitivity functions suffered with the existence of false-locks, and do not happen to exhibit global convergence.

During the same period, some researchers investigated decision-free sensitivity functions admissible enough to ensure global convergence. An admissible decision-free update is characterized by the following update:

$$\phi_{k+1} = \phi_k + \mu\big(z_{k,I}^2 - z_{k,R}^2\big) z_{k,I} z_{k,R} = \phi_k - \frac{1}{4}\mu\, \Im[z_k^4], \tag{10}$$

where $z_{k,I}$ and $z_{k,R}$ are the quadrature and in-phase components of z_k, respectively. It is the first globally convergent blind algorithm for carrier-phase recovery in QAM systems. We refer to it as (adaptive) fourth-power phase recovery algorithm (FP-PRA).

Historically, the update (10) was first proposed by Vandamme et al. in the year 1979 [16]. In 1980, the feasibility of fourth-order statistics for phase recovery in QAM was addressed by Frank [17]. However, these two works failed to attract attention from adaptive filtering and communication communities until it appeared in a monograph by Benveniste and Metivier [18]. During 1990's, under the framework of adaptive blind source separation, the same algorithm was rediscovered independently by a number of researchers including Macchi and Moreau [19,20], Mathis [11], Belouchrani and Ren [21], and Zarzoso and Nandi [22]. During the same period, under the context of carrier phase recovery, the fourth-order statistics appeared as an adaptive solution for phase recovery in [7,23–25]. More recent work (during 2000–2010) can be found in [26–32] where similar fourth-order statistics are presented for adaptive phase recovery in VSB systems. As far as the analysis is concerned, the link between the steady-state estimates of FP-PRA and decision-directed loop is established in [18]. In [33], authors discussed the second-order realization of FP-PRA. In the year 2010, the mean-time-to-loose-lock of FP-PRA is discussed in [34] using a rigorous stochastic differential equation framework. The latest appearance is [35] where the update (10) is used to track phase noise in coherent optical QAM-based OFDM systems.

The update (10) results by the minimization of $J(z_k) = \min_{\phi_k} \mathbb{E}\big|z_k^4 + 1\big|^2$ [18].[5] In [36], it was shown that FP-PRA can be obtained by minimizing a relatively simpler cost[6]

$$J(z_k) = \min_{\phi_k} \mathbb{E}\left(z_{k,R}^4 + z_{k,I}^4\right), \tag{11}$$

while in [37], Abrar showed that FP-PRA can be obtained equivalently from the following cost function[7]:

$$J(z_k) = \min_{\phi_k} \mathbb{E}\left(z_{k,R}^2 - z_{k,I}^2\right)^2. \tag{12}$$

Ignoring noise, (12) leads to

[4] For a complex-valued z, $\mathrm{csign}[z] = \mathrm{sign}[\Re[z]] + \mathbf{i}\,\mathrm{sign}[\Im[z]]$.

[5] Exploiting the fact $\frac{\partial}{\partial\phi_k} z_k^n = -\mathbf{i} n z_k^n$ for $n \in \mathbb{Z}$, we obtain

$$\begin{aligned}
\frac{\partial}{\partial\phi_k} J(z_k) &= \frac{\partial}{\partial\phi_k}\big|z_k^4 + 1\big|^2 = \frac{\partial}{\partial\phi_k}\big(z_k^4 + 1\big)\big(z_k^4 + 1\big)^* \\
&= \big(z_k^4 + 1\big)^* \frac{\partial}{\partial\phi_k}\big(z_k^4 + 1\big) + \big(z_k^4 + 1\big)\frac{\partial}{\partial\phi_k}\big(z_k^4 + 1\big)^* \\
&= \big(z_k^4 + 1\big)^*\big(-\mathbf{i}4 z_k^4\big) + \big(z_k^4 + 1\big)\big(\mathbf{i}4 (z_k^4)^*\big) \\
&= 8\left(\frac{z_k^4 - (z_k^4)^*}{\mathbf{i}2}\right) = 8\,\Im\left[z_k^4\right] = 32(z_{k,R}^2 - z_{k,I}^2) z_{k,I} z_{k,R}
\end{aligned}$$

Merging the coefficient with μ, and substituting the gradient in $\phi_{k+1} = \phi_k - \mu\frac{\partial}{\partial\phi_k} J(z_k)$, we obtain (10).

[6] The (stochastic approximate) gradient descent update is given by $\phi_{k+1} = \phi_k - \mu\frac{\partial}{\partial\phi_k}\big(z_{k,R}^4 + z_{k,I}^4\big)$. Exploiting the facts: $\frac{\partial}{\partial\phi_k} z_{k,R} = +z_{k,I}$ and $\frac{\partial}{\partial\phi_k} z_{k,I} = -z_{k,R}$, the update (10) is immediately obtained.

[7] Integrating the additive inverse of sensitivity function with respect to ϕ_k gives the cost function:

$$\begin{aligned}
&\int (z_{k,R}^2 - z_{k,I}^2) z_{k,I} z_{k,R}\, \mathrm{d}\phi_k \\
&= \frac{1}{4}\int (z_{k,R}^2 - z_{k,I}^2)(4 z_{k,I} z_{k,R})\, \mathrm{d}\phi_k \\
&= \frac{1}{4}\int (z_{k,R}^2 - z_{k,I}^2)\, \mathrm{d}(z_{k,R}^2 - z_{k,I}^2) \\
&= \frac{1}{8}(z_{k,R}^2 - z_{k,I}^2)^2 + \text{constant}.
\end{aligned}$$

$$J(\psi) = \mathbb{E}\left(z_{k,R}^2 - z_{k,I}^2\right)^2 = \mathbb{E}\big[0.5|x|^4 - \big(3x_I^2x_R^2 - \tfrac{1}{2}x_R^4 - \tfrac{1}{2}x_I^4\big)\cos(4\psi)\big]. \quad (13)$$

The first and second derivative of $J(\psi)$ are given as

$$J^{(1)}(\psi) = 4\mathbb{E}\big[3x_I^2x_R^2 - \tfrac{1}{2}x_R^4 - \tfrac{1}{2}x_I^4\big]\sin(4\psi),$$
$$J^{(2)}(\psi) = 16\mathbb{E}\big[3x_I^2x_R^2 - \tfrac{1}{2}x_R^4 - \tfrac{1}{2}x_I^4\big]\cos(4\psi). \quad (14)$$

Observe that $J(\psi)$ has the following stationary points:

$$\text{Local Minima: } \psi = 0, \pm\pi/2, \pm\pi, \cdots$$
$$\text{Local Maxima: } \psi = \pm\pi/4, \pm 3\pi/4, \pm 5\pi/4, \cdots \quad (15)$$

Observe that $J(\psi)$ has the same number of local minima and maxima and no saddle points. It yields a desirable local minimum at $\psi = 0$, but with a possible ambiguity of 90°. Also notice that, as required, the estimator is unimodal in $-\pi/4 < \psi \le \pi/4$. Notice that, however, it requires $\mathbb{E}[3x_I^2x_R^2 - x_R^4] =: -\kappa > 0$, which is true for practical QAM sources due to their sub-Gaussian nature.[8]

With noise, $z_k = x_k \exp(\mathbf{i}\psi_k) + e_k$, we have $J = \mathbb{E}[\frac{1}{2}x_I^4(\cos(4\psi)+1) + \frac{1}{2}x_R^4\cos(4\psi) + e_I^4 + e_R^4 + x_I^2(4e_I^2\cos(2\psi) - 4e_R^2\cos(2\psi) - 3x_R^2\cos(4\psi) + 2e_I^2 + 2e_R^2 + x_R^2) + \frac{1}{2}x_R^4 - 2e_I^2e_R^2 + 2e_I^2x_R^2 + 2e_R^2x_R^2 - 4e_I^2x_R^2\cos(2\psi) + 4e_R^2x_R^2\cos(2\psi)]$. For $\mathbb{E}e_R^2 = \mathbb{E}e_I^2$, we obtain $J = \mathbb{E}[\frac{1}{2}x_I^4(\cos(4\psi)+1) + \frac{1}{2}x_R^4\cos(4\psi) + e_I^4 + e_R^4 + x_I^2(x_R^2 - 3x_R^2\cos(4\psi) + 2e_I^2 + 2e_R^2) + \frac{1}{2}x_R^4 - 2e_I^2e_R^2 + 2e_I^2x_R^2 + 2e_R^2x_R^2]$. Exploiting $\mathbb{E}e_I^2 = \mathbb{E}e_R^2 = \sigma_e^2$ and $\mathbb{E}e_I^4 = \mathbb{E}e_R^4 = 3\sigma_e^4$, and simplifying we obtain

$$J(\psi) = (\mathbb{E}x_R^2)^2 + \mathbb{E}x_R^4 - \big\{3(\mathbb{E}x_R^2)^2 - \mathbb{E}x_R^4\big\}\cos(4\psi) + 4\sigma_e^2\big(\sigma_e^2 + 2\mathbb{E}x_R^2\big) \quad (16)$$

which clearly indicates that the locations of minima/maxima are unaffected with the presence of noise. So, even in noisy environment, the FP-PRA exhibits globally convergent property.

4. Modified phase recovery algorithms

In this section, two methods are suggested and implemented to enhance the performance of basic FP-PRA synchronizer. These methods rely largely on the observation that, in the absence of phase-offset, the (complex-valued baseband) data symbol residing on the $\pi/4$-diagonal contribute zero update.

In the first method, which will be termed as *QPSK-like partitioning of QAM*, we consider fixing phase-offset in only those QAM symbols which reside originally on the diagonal and for which no other non-diagonal symbol exhibits the similar modulus. Like in 64-QAM, the data symbols $\pm 1 \pm \mathbf{i}1$, $\pm 3 \pm \mathbf{i}3$, and $\pm 7 \pm \mathbf{i}7$ reside on the $\pi/4$-diagonal and no other non-diagonal data symbol exhibit similar modulus; they are QPSK-like. On the other hand, say in the first quadrant, the modulus of $5 + \mathbf{i}5$ is similar to those of two other non-diagonal symbols $1 + \mathbf{i}7$ and $7 + \mathbf{i}1$; so $5 + \mathbf{i}5$ is not QPSK-like. Therefore, under the assumption of perfect gain control, any phase rotation in QPSK-like data symbols can be detected by measuring their positions with respect to their respective diagonals, and corrected accordingly.

[8] For square-QAM, we have $\mathbb{E}x_I^2x_R^2 = \mathbb{E}x_I^2\mathbb{E}x_R^2$ (independency), and $\mathbb{E}x_I^2 = \mathbb{E}x_R^2$ (symmetry). Based on which $\kappa = \mathbb{E}x_R^4 - 3(\mathbb{E}x_R^2)^2$ becomes a measure of (unnormalized) kurtosis of the in-phase component, and which is negative due to uniform (sub-Gaussian) distribution of x_R (and this is also true for x_I). A similar discussion for cross-QAM is not straightforward due to lack of the independency. Numerically, however, we can show that the value of κ is negative for all practical QAM like it is equal to −34, −113, −546, −1810, and −8738 for 16-, 32-, 64-, 128-, and 256-QAM, respectively. If κ is normalized by $(\mathbb{E}x_R^2)^2$, then we obtain $\kappa/(\mathbb{E}x_R^2)^2$ to be equal to −1.3600, −1.2381, −1.2094, −1.2023 for 16-, 64-, 256-, and 1024-QAM, respectively, which can be seen to approach −1.2000 – a theoretical limit of kurtosis for uniform distribution.

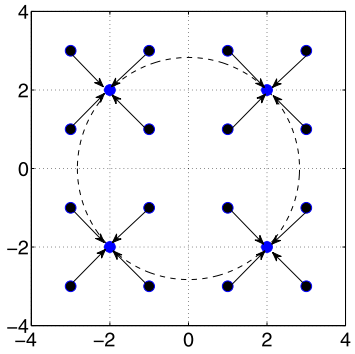

Fig. 3. Transformation of 16-QAM to QPSK.

In the second method, the QAM symbols are transformed to correspond a QPSK constellation. Consider a 16-QAM constellation, where coordinates are $x_k \in \{\pm 1 \pm \mathbf{i}, \pm 3 \pm \mathbf{i}, \pm 1 \pm \mathbf{i}3, \pm 3 \pm \mathbf{i}3\}$. The following QAM-to-QPSK transformation is suggested in [38] for 16-QAM:

$$s_k = \big[x_{k,R} - \text{sign}\big(x_{k,R} - 2\text{sign}(x_{k,R})\big)\big] + \mathbf{i}\big[x_{k,I} - \text{sign}\big(x_{k,I} - 2\text{sign}(x_{k,I})\big)\big] \quad (17)$$

where sign is standard signum function. The four transformed symbols (s_k) now appear on the (dotted) constant modulus as depicted in Fig. 3. In the sequel, we discuss in detail the feasibility of such transformation for square/cross-QAM in the presence of phase imperfections.

4.1. QPSK-like partitioning of QAM

QPSK-like partitioning of QAM constellation has been exploited in a number of recent articles including [38–49] and they have been studied largely for 16- and/or 64-QAM. Historically, the idea of QPSK partitioning of QAM signal first appeared in the work of Horikawa and Saitos in the year 1980 [50]. According to which only a subset of diagonal QAM symbols (which resemble QPSK) are used for phase/frequency-offsets recovery and the rest of the non-diagonal symbols are discarded.

In the year 1988, Sari and Moridi [15] revisited the Horikawa–Saitos's idea and described the partitioning based stop-and-go scheme in the following words: "some windows are drawn around the diagonal points, and the measured phase error is fed to the loop filter only when the received signal point falls within one of these windows. The previous phase detector output is held in memory and used as the current input to the loop filter otherwise". So they suggested to make windows (depicted as square decision boxes in Fig. 4) representing *reliable* zones around *QPSK-like* diagonal symbols and they used the following expression as a measure to obtain an estimate of phase-offset:

$$z_{k,I}\,\text{sign}(z_{k,R} - \widehat{a}_{k,R}) - z_{k,R}\,\text{sign}(z_{k,I} - \widehat{a}_{k,I}), \quad (18)$$

where $\widehat{a}_{k,R}$ and $\widehat{a}_{k,I}$ are the decisions made on $z_{k,R}$ and $z_{k,I}$, respectively, only when the symbol z_k falls in the square decision box.

4.1.1. Proposed modifications

Here we suggest two modifications in the work of Sari and Moridi. First, instead of *square* windows, we suggest *annular* regions to be used as reliable zones where phase recovery may be carried out.[9] Secondly, unlike (18), we suggest updates which are not dependent on decision symbols.

[9] The use of annular regions has appeared in [7] under the context of maximum likelihood based estimator for phase recovery so that the symbols falling in specific

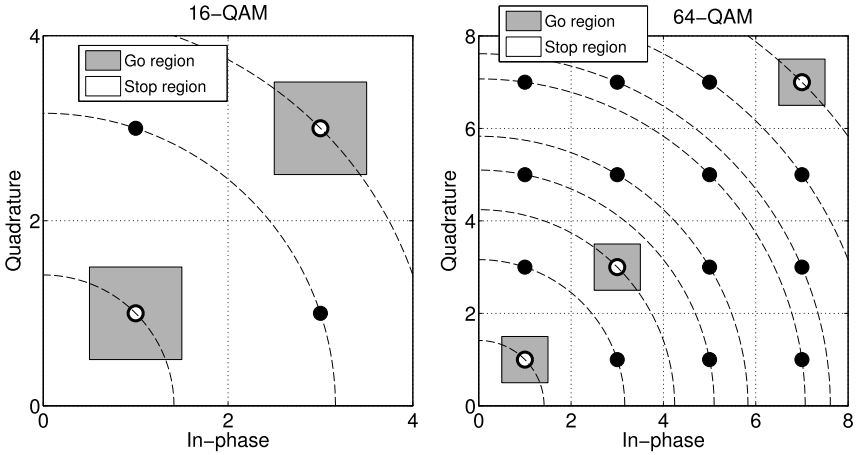

Fig. 4. Sari-Moridi stop-and-go regions for (a) 16-QAM, (b) 64-QAM.

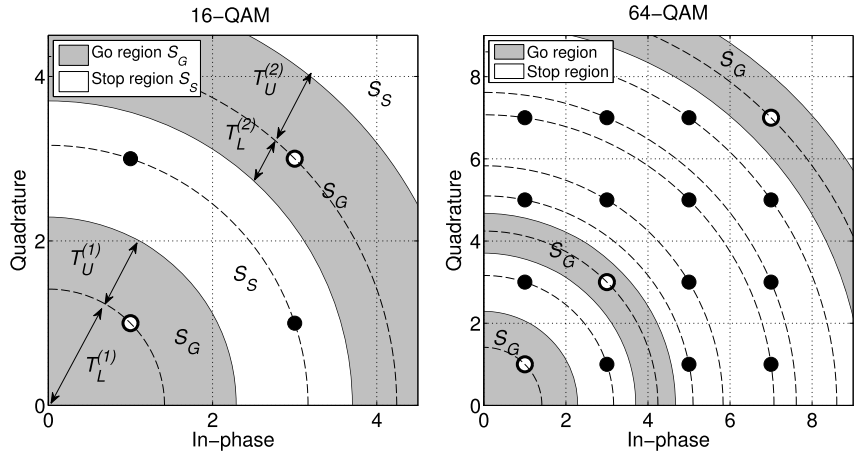

Fig. 5. Stop-and-go regions for (a) 16-QAM, (b) 64-QAM.

Refer to Fig. 5 for QPSK partitioning of two square-QAM signals where only one quarter of the signal constellation is shown for convenience. Symbols depicted as filled markers (●) are those which are not used in phase recovery process while a subset of symbols depicted as unfilled markers (∘) are those *QPSK-like* symbols which are used in phase recovery process. The grey annuli regions (toroidal regions) are those *windows* where the update is allowed and we label them as *Go* regions while the rest of the space (the white space) is *Stop* region where the update is not allowed. Note also that not all diagonal symbols are qualified to be designated as *QPSK-like* symbols rather only those which do not share their respective annuli (with any other distortion-free symbols) are classified as *QPSK-like*.

Let S_G and S_S denote the *Go* and *Stop* regions, respectively. Referring to Fig. 5(a), the thresholds ($T_l^{(i)}$ and $T_u^{(i)}$) which are used to decide annuli (S_G) can be set to some optimum values which may depend on the operating conditions [39]. In our simulations, however, we have obtained pair of thresholds for each *QPSK-like* symbol by simply bi-partitioning the annuli. This stop-and-go strategy, however, strictly requires the recovery of true energy of the signal constellation in order to distinguish between signal points with amplitude variation. In practice, this can be achieved by using automatic gain control and digital signal processing methods [39].

The diagonal d_k can be related to z_k as follows:

$$d_k = \operatorname{sign}(z_{k,R}) + \mathbf{i}\operatorname{sign}(z_{k,I}) =: \operatorname{csign}(z_k). \tag{19}$$

Given $z_k = y_k \exp(-\mathbf{i}\phi_k)$, let $z_k \in S_G$ lie in the first quadrant. The angle between $z_k = |z_k|\exp(\mathbf{i}(\psi_k + \pi/4))$ and the diagonal d_k is the residual phase-offset ψ_k, where ψ_k could be positive- or negative-valued depending on whether z_k lies above or below the diagonal line, respectively. So an estimate of ψ_k can be obtained as

$$z_k \operatorname{csign}(z_k^*) = \sqrt{2}\,|z_k| \exp(\mathbf{i}\,\psi_k). \tag{20}$$

Thus, given z_k, an estimate of ψ_k can be obtained as follows:

$$\psi_k = \sin^{-1}\left(\frac{\Im\left[z_k \operatorname{csign}(z_k^*)\right]}{\sqrt{2}\,|z_k|}\right) = \cos^{-1}\left(\frac{\Re\left[z_k \operatorname{csign}(z_k^*)\right]}{\sqrt{2}\,|z_k|}\right), \tag{21}$$

where $\Re[z_k \operatorname{csign}(z_k^*)] = |z_{k,R}| + |z_{k,I}|$ and $\Im[z_k \operatorname{csign}(z_k^*)] = z_{k,I}\operatorname{sign}(z_{k,R}) - z_{k,R}\operatorname{sign}(z_{k,I})$. Assuming a small ψ, i.e., $\sin(\psi) \approx \psi$, and ignoring the denominator, we obtain *stop-and-go diagonal-based phase recovery algorithm (SGD-PRA)* as follows:

SGD-PRA:

$$\phi_{k+1} = \begin{cases} \phi_k + \mu\Im\left[z_k \operatorname{csign}(z_k^*)\right], \\ \quad \text{if } z_k \in S_G \text{ (Go)}. \\ \phi_k, \ \text{if } z_k \notin S_G \text{ (Stop)}, \end{cases} \tag{22}$$

where the factor $\sqrt{2}$ has been merged with the step-size μ. Another possibility is to exploit stop-and-go strategy in FP-PRA lead-

annular regions were used and others were discarded. Here, however, these regions are used to devise a stop-and-go strategy for an adaptive phase-recovery scheme.

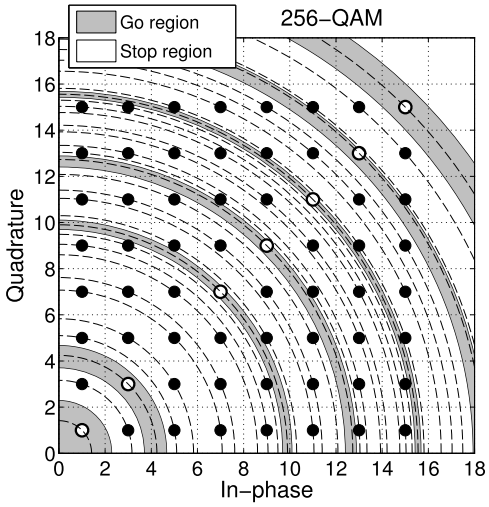

Fig. 6. Stop-and-go regions in 256-QAM constellation.

ing to *stop-and-go fourth-power phase recovery algorithm (SGFP-PRA)* as follows:

$$\text{SGFP-PRA:}\quad \phi_{k+1} = \begin{cases} \phi_k + \mu(z_{k,I}^2 - z_{k,R}^2) z_{k,I} z_{k,R}, \\ \qquad \text{if } z_k \in S_G \text{ (Go)}. \\ \phi_k, \ \text{if } z_k \notin S_G \text{ (Stop)}. \end{cases} \tag{23}$$

4.1.2. Remarks

1. Unlike [38–42,44,47], where the authors have exploited different variants of (maximum-likelihood-based [51]) feed-forward estimators, our proposed algorithms SGD-PRA and SGFP-PRA are adaptive (feedback) in nature. Moreover, unlike [50] and [15], the proposed algorithms do not rely on decision symbols and exploit solely the derotated symbol z_k.
2. It can be noted that the sensitivity of SGD-PRA is proportional to $\sin(\psi)$ for $|\psi| \leq \pi/4$. On the other hand, the sensitivity of SGFP-PRA is proportional to $\sin(4\psi)$ for $|\psi| \leq \pi/4$. So the sensitivity of SGD-PRA is larger than SGFP-PRA, and we expect to notice faster convergence to be exhibited by SGD-PRA than SGFP-PRA.
3. For QAM constellations, larger than 64, there exist very few *QPSK-like* symbols and thus very little reliable region is available to facilitate the recovery of phase. Consider the 256-QAM signal in Fig. 6. The situation becomes even worse when the SNR is not high enough. A possible solution for higher order constellations is to somehow increase the reliable regions or increase the number of reliable symbols. In the next section, we discuss how to *make all QAM symbols transformed into QPSK-like symbols yielding an always-go phase recovery mechanism.*

4.2. QAM-to-QPSK transformation

In the year 2012, Cartledge and his colleagues came up with a novel idea to transform a 16-QAM constellation into a QPSK constellation [38,44]. Just recently, the idea has been extended to 32-/128- (cross) and 64-QAM (higher-order square) in [52] and [53], respectively. The main idea was to transform QAM alphabets into QPSK alphabets such that, in the absence of any phase/frequency-offset, the true alphabets of each quadrant gets mapped nearly in the middle and on to the diagonals of their respective quadrants. In the presence of any phase imperfection, however, it has been shown numerically that the (majority of) transformed alphabets of QAM lie above and below the diagonal for positive and negative values of phase-offset, respectively (say for first quadrant). The idea was heuristically presented and experimentally (as well as numerically) justified to be effective. In this work, we examine the idea analytically and discuss its feasibility for both square- and cross-QAM.

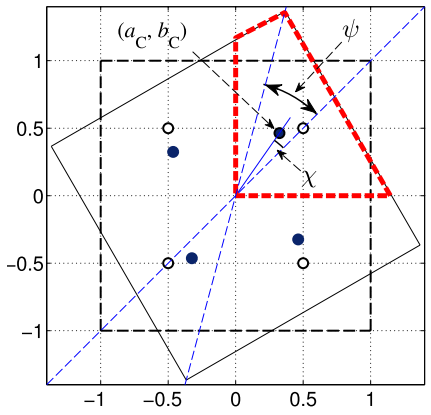

Fig. 7. The square disc (dotted line) and its rotated copy (solid line). The angle of rotation is ψ (anti-clockwise), and the centroid moves by the angle χ. The marker ○ denotes positions of quadrant-wise centroids during correct orientation, while marker • denotes the positions after rotation.

In the physical sense, the transformed four complex-valued QAM-to-QPSK symbols can be understood as the *centroids* (centers of masses) of signal constellation in each of the four quadrants. Under no phase-offset, the number of QAM symbols (in all four quadrants) above and below the diagonal is same, it makes the centroid of each quadrant lie onto the *diagonal*, and the synchronizer is ideally required to make no update. However, when the constellation is subjected to rotation, due to non-circular geometry of QAM (and specifically due to the corner points), the centroid of each quadrant moves in the direction of rotation.

4.2.1. Square-QAM

To quantify the movement of centroid of a square-QAM, consider a uniform square (disc), lying on $[-1,1] \times [-1,1]$, which is subjected to rotation with phase residual $\psi \in (-\pi/4, \pi/4]$ (refer to Fig. 7). Let (a_C, b_C) denote the coordinates of centroid of the rotated square in the first quadrant. It is noticed numerically that the clockwise or anticlockwise rotation in the square makes the centroid (say in the first quadrant) move above and below the diagonal, respectively. If $\psi \neq 0, \pm\pi/4$ then centroid is at an angle χ away from the diagonal. It is easy to show that (see Appendix A)

$$\chi = \tan^{-1}\left(\tan(\psi)\frac{2\cos(\psi)^2 - 1}{2\cos(\psi)^2 + 1}\right). \tag{24}$$

So the movement of the location of centroid captures the amount of rotation. Refer to Fig. 8(a) where the loci of the centroid of a uniform square is depicted for the first quadrant when ψ changes from $-\pi/4$ to $\pi/4$. Further, using Taylor's series expansion of (24) in Mathematica®, we can show that $\chi = \frac{1}{3}\psi - \frac{28}{81}\psi^3 - \frac{52}{243}\psi^5 - \frac{1352}{10,935}\psi^7 + O[\psi^9]$. For small ψ, we can see that $\chi \approx \psi/3$; so χ is related directly with ψ, and can be used as an admissible estimate of ψ. In order to compare the sensitivities of fourth-power (FP) and centroid-steering (CS) methods, we expand (24) as a factored power-series, we obtain

$$\chi \approx \left(\frac{1}{12} + \frac{4}{29}\psi^2 + \frac{1604}{12,615}\psi^4 + \cdots\right)\sin(4\psi). \tag{25}$$

Due to all positive higher-order terms, it is apparent from (25) that any synchronizer based on the information of centroid can prove to be more sensitive than the fourth-power method.

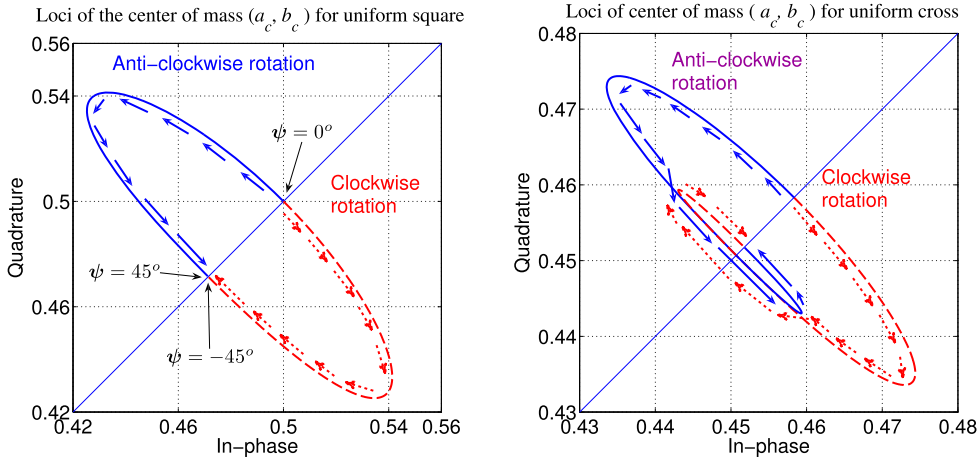

Fig. 8. (a) Loci of the coordinates (a_C, b_C), the centroid, as a function of phase error ψ in the first quadrant for a uniformly dense square. (b) Loci of the coordinates (a_C, b_C), centroid, as a function of phase error ψ in the first quadrant for a uniform density cross. False-lock can be observed to lie in the ranges $-45° < \psi \leq -33.69°$ and $33.69° \leq \psi \leq 45°$.

4.2.2. Cross-QAM

Now we consider a uniform cross (disc). Unlike a uniform square, the centroid of a uniform cross, however, is not always above and below the diagonal, respectively, for a clockwise or anticlockwise rotation. For a uniform cross, when subjected to a clockwise rotation, it can be noticed that the centroid is above the diagonal provided the phase error is less than $\tan^{-1}(2/3)$ radian. Beyond this value, in the range $\psi \in (\tan^{-1}(2/3), \pi/4]$, the centroid moves below the diagonal and creates a false-lock; this behavior is depicted in Fig. 8(b). Similarly, for an anticlockwise rotation, false-lock is found in the range $\psi \in (-\pi/4, -\tan^{-1}(2/3))$. It is important to understand that why the mapping-to-the-centroid fails to yield a good sensitivity function for a cross-QAM. To understand this phenomenon, we should look into the geometrical differences between a square- and cross-QAM; cross differs from square due to having no corners on the diagonals. For a uniform cross, before the false lock, we have (see Appendix B)

$$\tan(\chi) = \tan(\psi)\frac{13\cos(\psi)^2 - 9}{13\cos(\psi)^2 + 9}, \tag{26}$$

while during the false lock, we have (see Appendix B)

$$\tan(\chi) = \frac{\left\{\cot(\psi)\Big(30\sin(2\psi)^2 + 30\sin(2\psi) - 120\sin(\psi)^2 - 23\sin(4\psi)\Big)\right\}}{30\sin(2\psi)^2 + 62\sin(2\psi) - 23\sin(4\psi)}. \tag{27}$$

The relation between χ and ψ is depicted in Fig. 9 for both uniform square and cross, where the sensitivity of the fourth-power ($\mathcal{S}(\psi) = \sin(4\psi)$ with appropriate scaling factor) is also plotted for the sake of comparison. Clearly the use of the knowledge of χ for the estimation of unknown phase-offset residual (ψ) is likely to be admissible for square-QAM. For a uniform cross, however, there exist possible false locks; more discussion on this issue for practical cross-QAM is provided in Section 8.1.

4.2.3. Obtaining centroid

In a block processing scenario, the computation of centroid is simple; symbol points in quadrants two, three and four are first needed to get mapped in quadrant one. It can easily be done by swapping the abscissa with ordinate (or not), with proper change in polarities. Then, by computing the mean of the symbol points, we can locate the centroid and determine the resulting angle χ. However, in an adaptive scenario, where we need to process on

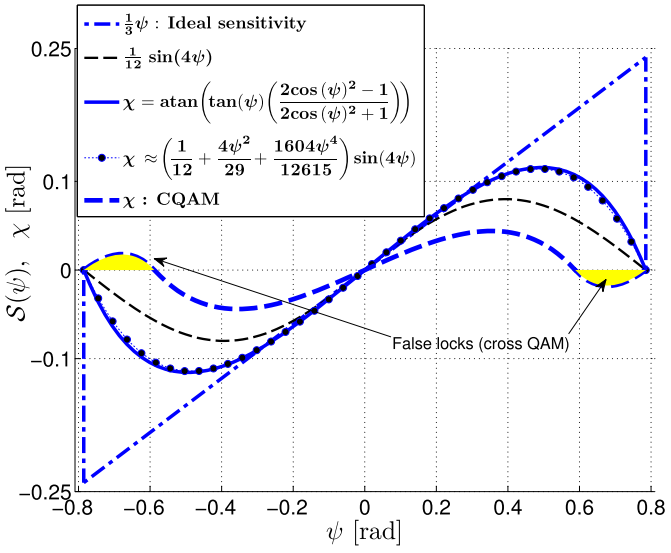

Fig. 9. Relation between ψ and χ.

symbol-by-symbol basis, alternative methods are required to directly compute the position of centroid. In [38,44], authors have suggested a QAM-to-QPSK mapping method for 16-QAM where the transformed QPSK is a good representation of the centroid. In this work, we generalize the existing art for any order of square- or cross-QAM (and which, to the best of our knowledge, is neither available nor properly discussed in open literature).

For an M-ary constellation, we have p unique positive levels for each of the quadrature components,

$$p = \begin{cases} \frac{\sqrt{M}}{2}, & \text{(Square-QAM)} \\ \frac{3\sqrt{2M}}{8}, & \text{(Cross-QAM)} \end{cases} \tag{28}$$

Note that 1 and $(2p-1)$, respectively, are the smallest and the largest positive levels. The center of the correctly-oriented signal constellation in the first quadrant (as an estimate for the position of centroid) can be obtained as $(\frac{q+1}{2}, \frac{q+1}{2}) = (p, p)$, where $q := (2p-1)$. Since the quadrature amplitude levels of adjacent alphabets differ by 2 units, we require at most $(p-1)$ iterations to drag the farthest point to the center (at p) where each itera-

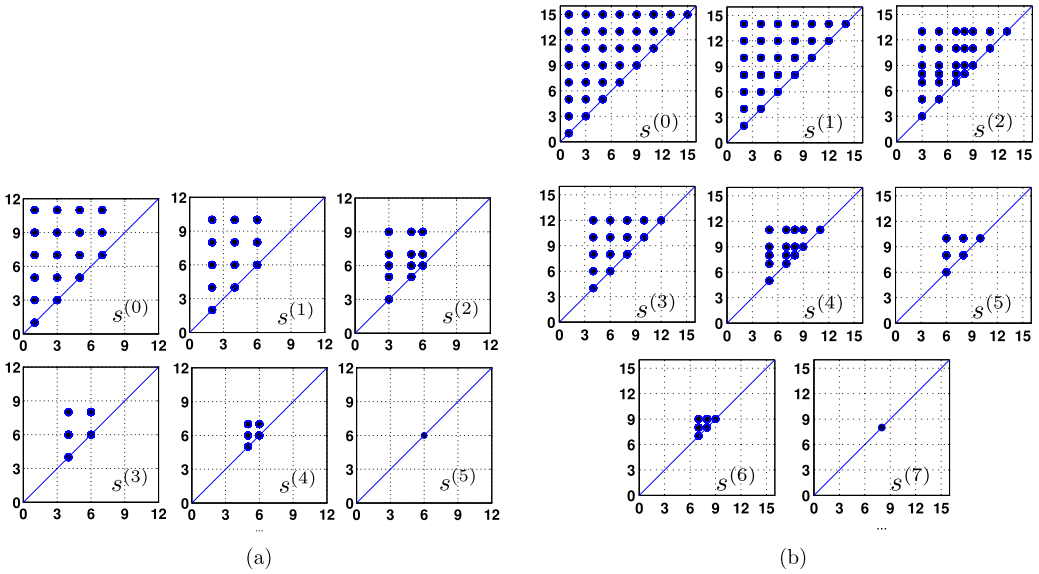

Fig. 10. Mapping to the center of mass in a no-noise and no-rotation scenario. (a) 128-QAM, and (b) 256-QAM.

tion reduces the distance by 1 unit. The QAM-to-QPSK mapping algorithm is summarized below:

$$
\begin{aligned}
&\widetilde{z}_k = \text{qam2qpsk}(z_k, p)\\
&s_R^{(0)} \leftarrow \Re[z_k]\\
&s_I^{(0)} \leftarrow \Im[z_k]\\
&\textbf{for } n = 1 : p-1 \textbf{ do}\\
&\quad s_R^{(n)} \leftarrow s_R^{(n-1)} - \text{sign}\big(s_R^{(n-1)} - p \cdot \text{sign}\big(s_R^{(n-1)}\big)\big)\\
&\quad s_I^{(n)} \leftarrow s_I^{(n-1)} - \text{sign}\big(s_I^{(n-1)} - p \cdot \text{sign}\big(s_I^{(n-1)}\big)\big)\\
&\textbf{end for}\\
&\widetilde{z}_k \leftarrow \frac{s_R^{(p-1)} + \mathbf{i}\, s_I^{(p-1)}}{p}\\
&\textbf{return}
\end{aligned}
\tag{29}
$$

The algorithm (29) is not fully optimized in the sense that it iterates for all $(p-1)$ times even if the symbol lies right at the middle; however, it ensures admissibly that a symbol, say residing above the diagonal (in a correctly oriented constellation), does not go below the diagonal during intermediate steps of transformation, and vice versa. This is demonstrated in Fig. 10(a) and (b) for 128- and 256-QAM, respectively, where symbols only in the upper-half of the constellation in the first quadrant are depicted for the purpose of clarity. It can be observed clearly that transforming symbols not only maintain the geometrical shape of the constellation during the course of transformation but (as required) reside in their upper-half region. The effect of noise and rotation on the performance of algorithm (29) is difficult to analyze analytically due to its iterative form and highly nonlinear structure. Empirically, however, we can examine its behavior in the presence of noise and rotation for different QAM sizes.

Under the presence of non-zero phase-offset ψ, we expect that the algorithm maps all (or majority of) symbols above or below the diagonal for anti-clockwise or clockwise rotation, respectively. Using the transformation method, we thus suggest an adaptive synchronizer, which we term as centroid-steering fourth-power phase recovery algorithm (CSFP-PRA), as given by:

$$
\begin{aligned}
&\text{CSFP-PRA:}\\
&\widetilde{z}_k = \text{qam2qpsk}(z_k, p)\\
&\phi_{k+1} = \phi_k - \frac{1}{4}\mu_k \Im[\widetilde{z}_k^4]
\end{aligned}
\tag{30}
$$

CSFP-PRA is similar to FP-PRA except that it uses transformed symbols to estimate the residual offset ψ.

5. Performance analysis

A good measure for the performance of synchronizer is the steady-state deviation (or variance). We define *mean square deviation* as $\text{MSD} =: \sigma^2 = \mathbb{E}\psi_\infty^2$ (rad^2) and compute it for (10). Let $\psi_k = \theta - \phi_k$ be the parameter error at time instant k, we have

$$\psi_{k+1} = \theta - \phi_{k+1} = \psi_k - \mu \mathcal{S}_k, \tag{31}$$

where $\mathcal{S}_k$ is some generic sensitivity function. Squaring and averaging the update, we get $\mathbb{E}\psi_{k+1}^2 = \mathbb{E}\psi_k^2 + \mu^2 \mathbb{E}\mathcal{S}_k^2 - 2\mu \mathbb{E}\psi_k \mathcal{S}_k$. During steady state, as $k \to \infty$, note that $\mathbb{E}\psi_{k+1}^2 = \mathbb{E}\psi_k^2$; this yields the following expression to solve for the expression for MSD (for any given phase recovery algorithm):

$$\lim_{k\to\infty} \mu \mathbb{E}\mathcal{S}_k^2 = \lim_{k\to\infty} 2\mathbb{E}\psi_k \mathcal{S}_k \tag{32}$$

which is true for any adaptive phase recovery algorithm with sensitivity function $\mathcal{S}_k$. It should be noted that we have not made any approximation in (32); it is an exact expression for the evaluation of MSD. There are two random variables (x_k and ψ_k) involved in (32), so we require conditional means for the evaluation of moments as given by: $\mathbb{E}\mathcal{S}_k^2 = \mathbb{E}_\psi[\mathbb{E}_x[\mathcal{S}_k^2 \mid \psi_k]] =: c_2 \mathbb{E}\psi_k^2 + c_3$, and $\mathbb{E}\psi_k \mathcal{S}_k = \mathbb{E}_\psi[\psi_k \mathbb{E}_x[\mathcal{S}_k \mid \psi_k]] =: \frac{c_1}{2}\mathbb{E}\psi_k^2$.

Evaluating these expectations, we obtain the following expression for its MSD:

$$\text{MSD} = \mathbb{E}\psi_\infty^2 = \frac{c_3 \mu}{c_1 - c_2 \mu} =: \sigma^2, \tag{33}$$

where the constants are obtained for FP-PRA as given by (see Appendix C):

$$c_1 := \mathbb{E}[12x_I^2 x_R^2 - 2x_I^4 - 2x_R^4], \tag{34a}$$

$$c_2 := \mathbb{E}[2x_R^8 - 56x_R^6 x_I^2 + 70x_R^4 x_I^4], \tag{34b}$$

$$\begin{aligned} c_3 := \mathbb{E}[&2x_I^2 x_R^6 - 2x_I^4 x_R^4 + 18\sigma_e^4 x_R^4 + 12\sigma_e^8 + 2\sigma_e^2 x_R^6\\ &+ 48\sigma_e^6 x_R^2 + 6\sigma_e^2 x_I^2 x_R^4 + 18\sigma_e^4 x_I^2 x_R^2]. \end{aligned} \tag{34c}$$

These values are applicable to SGFP-PRA and CSFP-PRA as well. In SGFP-PRA, however, we evaluate these constant considering only

those QAM symbols which are QPSK-like. Similarly, in CSFP-PRA, we consider only four transformed alphabets, that is, $\pm 1 \pm \mathbf{i}$.

The values of statistical constants c_i's for SGD-PRA are obtained as (see Appendix D):

$$c_1 := 4\mathbb{E}|x_R|, \tag{35a}$$

$$c_2 := 2\mathbb{E}|x_R x_I|, \tag{35b}$$

$$c_3 := \mathbb{E}|x|^2 + 2\sigma_e^2 - 2\mathbb{E}|x_R x_I|. \tag{35c}$$

By rearranging Equation (33), we can obtain the value of loop-gain for the required MSD and given signal statistics as given by $\mu = c_1\sigma^2/(c_3 + c_2\sigma^2)$.

Denote $\mathbb{E}\psi_k^2$ as the *instantaneous mean squared deviation* at time index k. Assume that the estimator is initialized with ϕ_0, which provides $\mathbb{E}\psi_0^2 = \mathbb{E}(\phi_0 - \theta)^2 = (\phi_0 - \theta)^2$. Next, substituting the values of moments $\mathbb{E}[\mathcal{S}_k^2|\psi_k]$ and $\mathbb{E}[\mathcal{S}_k|\psi_k]$ in (32), we obtain

$$\begin{aligned}
\mathbb{E}\psi_k^2 &= (1 - c_1\mu + c_2\mu^2)\mathbb{E}\psi_{k-1}^2 + c_3\mu^2 \\
\Rightarrow \mathbb{E}\psi_k^2 - \sigma^2 &= (1 - c_1\mu + c_2\mu^2)\mathbb{E}\psi_{k-1}^2 + c_3\mu^2 - \sigma^2 \\
&= (1 - c_1\mu + c_2\mu^2)\left(\mathbb{E}\psi_{k-1}^2 - \sigma^2\right) \\
&= (1 - c_1\mu + c_2\mu^2)^k\left(\mathbb{E}\psi_0^2 - \sigma^2\right).
\end{aligned} \tag{36}$$

Denoting and substituting $\rho = \mathbb{E}\psi_0^2$ and $\tau_\sigma = -\ln(1 - c_1\mu + c_2\mu^2)$, we obtain the dynamics of metric MSD as follows:

$$\mathbb{E}\psi_k^2 = \sigma^2 + \left(\rho - \sigma^2\right)\exp(-k\tau_\sigma). \tag{37}$$

Note that the parameter ρ is the initial value of MSD and τ determines the convergence speed of the adaptive synchronizer in the mean square error sense. Similarly, taking expectation of $\psi_k = \psi_{k-1} - \mu\mathcal{S}_{k-1}$, we can obtain $\mathbb{E}\psi_k = (\theta - \phi_0)\exp(-k\tau_\mu)$, where $\tau_\mu = -\ln\left(1 - \mu\frac{c_1}{2}\right)$ and it helps us to readily obtain

$$\mathbb{E}\phi_k = \theta\left(1 - \exp(-k\tau_\mu)\right) + \phi_0\exp(-k\tau_\mu) \tag{38}$$

which gives the dynamic convergence of ϕ_k.

Define the *convergence time* as the number of iterations, K, that is needed by $\mathbb{E}\psi_k^2$ to reach $(1 + \varepsilon)$ times its steady-state value $\sigma^2 = \mathbb{E}\psi_\infty^2$, for some $\varepsilon > 0$. That is, it is the time $k = K$ at which we have $\mathbb{E}\psi_K^2 = (1 + \varepsilon)\sigma^2$. At $k = K$, introducing (36), we get

$$K = \frac{\ln\left(\dfrac{\varepsilon c_3\mu}{(c_1 - c_2\mu)(\phi_0 - \theta)^2 - c_3\mu}\right)}{\ln(1 - c_1\mu + c_2\mu^2)} = \ln\left(\frac{\varepsilon\sigma^2}{\rho - \sigma^2}\right)^{-\frac{1}{\tau_\sigma}}. \tag{39}$$

5.1. Performance in the presence of frequency-offset

Due to the mismatch in the frequencies of local oscillators at the transmitter and the receiver ends, the resulting frequency-offset makes phase drift over time. Assume that the phase is drifting linearly at the (normalized) rate Ω, i.e., the true phase at the kth update is given by[10] $\phi_k = \theta + k\Omega$. The deviation of the estimated parameter ϕ_k from the true phase θ is thus given by: $\psi_k = \theta + k\Omega - \phi_k$, using which we obtain $\psi_{k+1} = \psi_k + \Omega - \mu\mathcal{S}_k$. Taking ensemble average, the steady-state system satisfies $\lim_{k\to\infty}\mathbb{E}\psi_{k+1} = \lim_{k\to\infty}\mathbb{E}\psi_k$. Under the assumption $\psi_k \ll 1$, we obtain the *tracking error* as given by $\mathbb{E}\psi_\infty \approx 2\Omega/(\mu c_1)$. In the presence of frequency-offset, the MSD can be obtained by combining the distortion contributed by acquisition and tracking phenomenon, which gives

$$\text{MSD} := \frac{c_3\mu}{c_1 - c_2\mu} + \frac{4\Omega^2}{\mu^2 c_1^2}\ (\text{rad}^2). \tag{40}$$

Notice that the first and the second term in (40) increases and decreases monotonically with the loop-gain μ, respectively. The optimal loop gain, μ^{opt} is obtained by seeking the minimum of (40). We need to solve $\partial\text{MSD}/\partial\mu = 0$, which gives $c_1^3c_3\mu^3 - 8c_2^2\Omega^2\mu^2 + 16c_1c_2\Omega^2\mu = 8c_1^2\Omega^2$. Solving for the positive (real-valued) root, we get the following optimum value of μ:

$$\mu^{\text{opt}} = \frac{(4c_4)^{\frac{1}{3}}}{3c_3c_1^3} - \frac{\Omega^2c_2 4^{\frac{5}{3}}\left(3c_3c_1^4 - 4\Omega^2c_2^3\right)}{3c_3c_1^3c_4^{\frac{1}{3}}} + \frac{8\Omega^2c_2^2}{3c_3c_1^3}, \tag{41}$$

where

$$\begin{aligned}
c_4 = &-144\Omega^4c_2^3c_3c_1^4 + 27c_3^2c_1^8\Omega^2 + 128\Omega^6c_2^6 \\
&+ 3^{\frac{3}{2}}\Omega^2c_3^2c_1^6\sqrt{\frac{27c_3c_1^4 - 32\Omega^2c_2^3}{c_3}}.
\end{aligned} \tag{42}$$

The optimum loop-gain depends on the statistics of the QAM signal and the frequency-offset Ω; note that, μ^{opt} increases with Ω.

5.2. Performance in the presence of phase noise

Consider a time-varying phase θ_k which varies according to the random-walk model

$$\theta_k = \theta_{k-1} + q_k, \quad \mathbb{E}q_k^2 = \sigma_q^2, \tag{43}$$

where q_k is a zero-mean Gaussian distributed i.i.d. sequence with variance σ_q^2. Assume further that θ_k is independent of $\{x_m\}$ and $\{y_m\}$ for all $m < k$. Denoting $\psi_k = \theta_k - \phi_k$, the following variance relation holds:

$$2\mathbb{E}\psi_k\mathcal{S}_k = \mu\mathbb{E}\mathcal{S}_k^2 + \mu^{-1}\sigma_q^2. \tag{44}$$

For proof, refer to [54] for the energy conservation based arguments for adaptive filter operating in the non-stationary environment. Considering the variance relation (44) and statistics c_i from (34a)–(34c), we have

$$\mathbb{E}\psi_\infty^2 = \frac{c_3\mu + \mu^{-1}\sigma_q^2}{c_1 - c_2\mu}. \tag{45}$$

The optimal step-size μ that tracks the phase noise is obtained by minimizing (45) with respect to μ, as given by

$$\mu^{\text{opt}} = \frac{\sqrt{c_3c_1^2\sigma_q^2 + c_2^2\sigma_q^4} - c_2\sigma_q^2}{c_1c_3} \approx \sqrt{\frac{1}{c_3}}\,\sigma_q, \tag{46}$$

yielding the minimum MSD

$$(\mathbb{E}\psi_\infty^2)_{\min} = \frac{c_3\mu^{\text{opt}} + \left(\mu^{\text{opt}}\right)^{-1}\sigma_q^2}{c_1 - c_2\mu^{\text{opt}}} \approx \frac{2\sigma_qc_3}{c_1\sqrt{c_3} - c_2\sigma_q} \approx \sqrt{\frac{4c_3}{c_1^2}}\,\sigma_q, \tag{47}$$

where approximations are based on the assumption that $c_3c_1^2 \gg c_2^2\sigma_q^2$. Note that both minimum MSD and optimal step-size are proportional to standard deviation of the phase noise, σ_q. So larger the deviation, the larger the optimum step-size to track with minimum MSD.

[10] The normalized frequency-offset Ω is defined as $\Omega = 2\pi f_\Delta/R_s$, where R_s is the symbol-rate and f_Δ is the difference between transmitter and receiver local-oscillator frequencies in Hertz.

5.3. Performance under both frequency-offset and phase noise

In the presence of both frequency-offset and phase noise with residual $\psi_k = \theta_k + k\Omega - \widehat{\theta}_k$, and assuming filter operation in steady-state, the MSD is given as

$$\text{MSD} = \frac{\mu\gamma + \mu^{-1}\sigma_q^2}{2\alpha - \mu\beta} + \frac{\Omega^2}{\mu^2\alpha^2}. \tag{48}$$

Minimization of which yields an optimal step-size:

$$\mu^{\text{opt}} = \frac{\zeta^2}{\sqrt[3]{\delta}} - \frac{\eta}{\sqrt[3]{\delta}} + \sqrt[3]{\delta} + \zeta, \tag{49}$$

where $\delta = \zeta^3 + \frac{2\Omega^2}{\alpha\gamma} - \frac{3}{2}\zeta\eta + \sqrt{\left(\zeta^3 + \frac{2\Omega^2}{\alpha\gamma} - \frac{3}{2}\zeta\eta\right)^2 - \left(\zeta^2 - \eta\right)^3}$, $\zeta = \frac{\Omega^2\beta^2 - \alpha^2\beta\sigma_q^2}{3\alpha^3\gamma}$, and $\eta = \frac{4\Omega^2\alpha\beta - \alpha^3\sigma_q^2}{3\alpha^3\gamma}$.

5.4. Performance with imperfect equalization

Now we discuss the effect of residual inter-symbol interference (ISI) on the performance of FP-PRA. We assume that the (equalized) input to the phase synchronizer is given as $y_k = \exp(\mathbf{i}\theta)\sum_{l=-\infty}^{\infty} h_l x_{k-l} + w_k$, where the elements in the sequence $\{h_l\}_{-\infty}^{\infty}$ are coefficients of overall channel-equalizer impulse response. Ignoring additive noise, the FP-PRA update may be expressed as:

$$\phi_{k+1} = \phi_k - \mu\Im\left[z_k^4\right] = \phi_k - \mu\Im\left[\left(y_k\exp(-\mathbf{i}\phi_k)\right)^4\right] \tag{50a}$$

$$\mathbb{E}\phi_{k+1} = \mathbb{E}\phi_k - \mu\Im\left[\mathbb{E}\exp(\mathbf{i}4\psi_k)\left(\textstyle\sum_l h_l x_{k-l}\right)^4\right] \tag{50b}$$

Under the assumptions that the data sequence $\{x_k\}$ is uncorrelated and zero-mean, we obtain (conditioned on ψ_k)

$$\begin{aligned}
&\mathbb{E}\exp(\mathbf{i}4\psi_k)\left(\textstyle\sum_l h_l x_{k-l}\right)^4 = \exp(\mathbf{i}4\psi_k)\mathbb{E}\left(\textstyle\sum_l h_l x_{k-l}\right)^4 \\
&= \exp(\mathbf{i}4\psi_k)\textstyle\sum_l\sum_m\sum_n\sum_r h_l h_m h_n h_r\,\mathbb{E}x_{k-l}x_{k-m}x_{k-n}x_{k-r} \\
&= \exp(\mathbf{i}4\psi_k)\left(\left(\mathbb{E}x_k^4\right)\textstyle\sum_l h_l^4 + 3\left(\mathbb{E}x_k^2\right)^2\left(\left(\sum_l h_l^2\right)^2 - \sum_l h_l^4\right)\right) \\
&= \left(\exp(\mathbf{i}4\psi_k)\textstyle\sum_l h_l^4\right)\mathbb{E}x_k^4,\ ^{11}
\end{aligned} \tag{51}$$

where we exploit the property $\mathbb{E}x_k^2 = 0$ owing to four-quadrant symmetry of QAM constellation. Further note that

$$\mathbb{E}x_k^4 = \underbrace{\mathbb{E}\left(x_{k,R}^4 + x_{k,I}^4 - 6x_{k,R}^2x_{k,I}^2\right)}_{<0} + \mathbf{i}4\underbrace{\mathbb{E}x_{k,R}x_{k,I}\left(x_{k,R}^2 - x_{k,I}^2\right)}_{=0}, \tag{53}$$

where the sub-Gaussian nature of QAM alphabets dictates $c := \mathbb{E}x_k^4 \in \mathbb{R}, c < 0$. Combining these findings, we obtain (below $c = -|c|$)

$$\lim_{k\to\infty}\left(\mathbb{E}\phi_{k+1} - \mathbb{E}\phi_k\right) = \mu\,|c|\,\Im\left[\exp(\mathbf{i}4\psi_k)\textstyle\sum_l h_l^4\right] = 0. \tag{54}$$

During steady-state, if the jitter due to non-zero step-size is ignored, then it is needed to ensure $\lim_{k\to\infty}\Im\left[\exp(\mathbf{i}4\psi_k)\sum_l h_l^4\right] \to 0$ in the vicinity of equilibrium point, which implies that $\left(4\psi_\infty + \arg\left(\sum_l h_l^4\right) \mod \frac{\pi}{2}\right) = 0$. These considerations lead to an interesting result as follows:

$$\phi_\infty = \theta + \frac{1}{4}\arg\left(\textstyle\sum_l h_l^4\right) \mod \frac{\pi}{2} \tag{55}$$

Clearly in the presence of ISI, the steady-state estimate may deviate from the target θ. It is mentioned in [18, p. 117] that the steady-state estimate of FP-PRA is insensitive to the additive noise and depends only on the residual ISI, so the above result is valid even with the consideration of AWGN noise. The proof is easy and the steps involve assuming that the AWGN is zero-mean, circular, zero-kurtosis, and uncorrelated with QAM alphabets.

Further assume that the residual ISI (after equalization process) is mild and satisfies the (open-eye) condition $h_\delta \gg \sum_{l\neq\delta}|h_l|$ and $h_\delta \in \mathbb{R}$ (where δ denotes the overall channel-equalizer processing delay). The update (55) can be expressed as:

$$\phi_\infty = \theta + \frac{1}{4}\tan^{-1}\frac{\Im\left[\sum_l h_l^4\right]}{\Re\left[\sum_l h_l^4\right]} \mod \frac{\pi}{2} \tag{56}$$

Note that $\Re\left[\sum_l h_l^4\right] \approx h_\delta^4$, and $\Im\left[\sum_l h_l^4\right] = \Im\left[\sum_{l\neq\delta} h_l^4\right]$. Further assuming small argument, $\tan^{-1}(z) \approx z$, and $|z| \ll \pi/4$, we get

$$\begin{aligned}
\tan^{-1}\frac{\Im\left[\sum_l h_l^4\right]}{\Re\left[\sum_l h_l^4\right]} \approx \tan^{-1}\frac{\Im\left[\sum_{l\neq\delta} h_l^4\right]}{h_\delta^4} &\approx \frac{\sum_{l\neq\delta}|h_l|^4\arg(4h_l)}{h_\delta^4} \\
&= 4\frac{\sum_{l\neq\delta}|h_l|^4\arg(h_l)}{h_\delta^4}.
\end{aligned} \tag{57}$$

Denoting $\omega := h_\delta^{-4}\sum_{l\neq\delta}|h_l|^4\arg(h_l)$, we summarize the finding as follows:

$$\phi_\infty \approx \theta + \omega \mod \frac{\pi}{2} \tag{58}$$

So the residual interfering paths ($l \neq \delta$) contribute an additional phase-offset, ω, which is directly proportional to their arguments (phases) and the fourth-power of their gains, and inversely proportional to the fourth-power of largest spike h_δ.

6. Adaptive estimation of the optimal step-size

In the last section, we evaluated closed-form expressions for the optimal step-sizes in the presence of frequency offset (FO) and phase noise (PN). However, it required a priori information of Ω and σ_q^2, which is not always known or may be changing over time. A possible solution is to find these optimal values blindly. The optimal value of step-size may be obtained adaptively by minimizing the fourth-power cost function (12) as the cost is sensitive to both FO and PN. The cost function is thus minimized (with respect to the variable step-size μ_k) only when the FO and PN are perfectly mitigated.

The step-size μ_k is known a priori to be *small* with high likelihood, therefore we may prefer to use sparse optimization method over traditional gradient based optimization [55]. Since it is also known that the step-size is *positive* real-valued, therefore, we can use Kivinen–Warmuth *natural* exponential gradient optimization method [56,57] for μ_k to minimizing the fourth-power phase-recovery cost function $\mathcal{J}(z_k)$ as follows:

$$\begin{aligned}
\mu_{k+1} &= \exp\left(\ln(\mu_k) - \frac{\alpha\mu_k}{4}\frac{\partial\mathcal{J}(z_k)}{\partial\mu_k}\right) \\
&= \mu_k\exp\left(-\frac{\alpha\mu_k}{4}\frac{\partial\mathcal{J}(z_k)}{\partial\mu_k}\right)
\end{aligned} \tag{59}$$

where $\mathcal{J}(z_k) := z_{k,R}^4 + z_{k,I}^4$, and α is a positive step-size. Using Taylor's series (first-order) expansion, we can obtain an approximate natural gradient version of (59) as follows:

$$\mu_{k+1} \approx \mu_k - \frac{\alpha}{4}(\mu_k)^2\frac{\partial\mathcal{J}(z_k)}{\partial\mu_k} \tag{60}$$

---

[11] $\mathbb{E}x_{k-l}x_{k-m}x_{k-n}x_{k-r}$

$$= \begin{cases} \mathbb{E}x_k^4, & \text{for } l=m=n=r. \\ \left(\mathbb{E}x_k^2\right)^2 & \text{for } l=m\neq n=r, \\ & \text{and } l=n\neq m=r, \\ & \text{and } l=r\neq m=n. \\ 0 & \text{otherwise.} \end{cases} \tag{52}$$

where

$$\frac{\partial \mathcal{J}(z_k)}{\partial \mu_k} = \frac{\partial \mathcal{J}(z_k)}{\partial \phi_k}\frac{\partial \phi_k}{\partial \mu_k} = \Im[z_k^4]\,\varpi_k, \tag{61}$$

$\varpi_k := \frac{\partial}{\partial \mu_k}\phi_k$, and $\Im[z_k^4] = 4\big(z_{k,R}^2 - z_{k,I}^2\big)z_{k,R}z_{k,I}$. Differentiating FP-PRA update with respect to μ_k gives an update for ϖ_k (below $\frac{\partial}{\partial \phi_k}\Im[z_k^4] = -4\Re[z_k^4]$)

$$\begin{aligned}\varpi_{k+1} &= \varpi_k - \frac{\mu_k}{4}\frac{\partial \Im[z_k^4]}{\partial \mu_k} - \frac{1}{4}\Im[z_k^4] \\ &= \varpi_k - \frac{\mu_k}{4}\frac{\partial \Im[z_k^4]}{\partial \phi_k}\frac{\partial \phi_k}{\partial \mu_k} - \frac{1}{4}\Im[z_k^4] \\ &= \varpi_k\Big(1 + \mu_k \Re[z_k^4]\Big) - \frac{1}{4}\Im[z_k^4].\end{aligned} \tag{62}$$

Finally, we obtain the following *variable step-size* variant of FP-PRA, that is VSSFP-PRA:

$$\boxed{\begin{aligned}&\text{VSSFP-PRA:}\\ &\phi_{k+1} = \phi_k - \frac{1}{4}\mu_k \Im[z_k^4]\\ &\varpi_{k+1} = \varpi_k\Big(1 + \mu_k \Re[z_k^4]\Big) - \frac{1}{4}\Im[z_k^4]\\ &\mu_{k+1} = \Big[\mu_k\Big(1 - \frac{\alpha}{4}\mu_k \varpi_k \Im[z_k^4]\Big)\Big]_{\mu_{\min}}^{\mu_{\max}}\end{aligned}} \tag{63}$$

where $\mu_{\min}$ and $\mu_{\max}$ are used appropriately for the purpose of squashing so that to avoid stuck-at-zero and divergence, respectively. This VSS method may easily be extended to other algorithms like CSFP-PRA. Care will be required as the step-size of CSFP-PRA happens to be relatively large. So to ensure stability, either we need to select very small value of α or limit (somehow) the magnitude of $\varpi_k\Im[z_k^4]$ which appears in the error term of μ_k update. We have found numerically that squashing the magnitude of $\varpi_k\Im[z_k^4]$ gives stable results. Thus, we suggest the following VSS variant of CSFP-PRA, that is VSSCSFP-PRA:

$$\boxed{\begin{aligned}&\text{VSSCSFP-PRA:}\\ &\widetilde{z}_k = \text{qam2qpsk}(z_k, p)\\ &\phi_{k+1} = \phi_k - \frac{1}{4}\mu_k \Im[\widetilde{z}_k^4]\\ &\varpi_{k+1} = \varpi_k\Big(1 + \mu_k \Re[\widetilde{z}_k^4]\Big) - \frac{1}{4}\Im[\widetilde{z}_k^4]\\ &\mu_{k+1} = \Big[\mu_k\Big(1 - \frac{\alpha}{4}\mu_k \,\text{sign}\big[\varpi_k \Im[\widetilde{z}_k^4]\big]\Big)\Big]_{\mu_{\min}}^{\mu_{\max}}\end{aligned}} \tag{64}$$

The aforementioned VSS algorithms are suitable for estimating optimal step-sizes in phase noise scenario. In the presence of frequency-offset, however, we would need a second-order algorithm, guidelines for which can be found in [7,8,58], and is not being described here due to the shortage of space.

7. Batch processing algorithms

Adaptive filtering is not always done on a symbol-by-symbol basis. In high-speed processing, we need to process several symbols in batch (or parallel) and the sequential algorithms need to be adjusted accordingly. The batch algorithms are known to be more robust to misadjustment which allows setting their step-sizes to higher values for faster convergence. Below we discuss how to realize the traditional FP-PRA in a batch manner by processing several received input symbols at a time.

Denoting $r_k = \exp(\mathbf{i}\phi_k)$, we have $z_k = r_k^* y_k$. So the phase synchronizer behaves like a one-tap equalizer. For the estimation of complex exponential r_k; consider the following problem:

$$\min_{r_k} \mathbb{E}\big|z_k^4 + 1\big|^2, \quad \text{subject to } |r_k| = 1 \tag{65}$$

By conditioning on the present value of r_k, considering $|r_k| = 1$, and exploiting the fact $\frac{\partial}{\partial r_k^*} r_k = 0$, we get

$$\begin{aligned}\frac{\partial}{\partial r_k^*}\mathbb{E}\big|z_k^4 + 1\big|^2 &= \frac{\partial}{\partial r_k^*}\mathbb{E}\Big(|y_k|^8 + (r_k^*)^4 y_k^4 + r_k^4 (y_k^*)^4 + 1\Big)\\ &= 4(r_k^*)^3 \mathbb{E} y_k^4,\end{aligned} \tag{66}$$

where the statistic $\mathbb{E}y_k^4$ may be computed from the last N_{b} number of received and stored symbols, and thus this computation will be required to be evaluated after every N_{b} number of received symbols. We suggest

$$\boxed{\mathbb{E}y_k^4 := \tfrac{1}{N_{\text{b}}}\textstyle\sum_{i=k-N_{\text{b}}+1}^{k} y_i^4} \tag{67}$$

where $k = N_{\text{b}}, 2N_{\text{b}}, 3N_{\text{b}}, \cdots$. These considerations lead to the following *forward driving iterative*[12] *batch processing* algorithm:

$$\boxed{\begin{aligned}&\textbf{for } k = N_{\text{b}} : N_{\text{b}} : \cdots \textbf{ do}\\ &\quad \xi \leftarrow \tfrac{1}{N_{\text{b}}}\textstyle\sum_{i=k-N_{\text{b}}+1}^{k} y_i^4\\ &\quad \textbf{for } n = 1 : N_{\text{i}} \textbf{ do}\\ &\quad\quad \widetilde{r}_n \leftarrow r_n - \mu\,(r_n^*)^3\,\xi\\ &\quad\quad r_{n+1} \leftarrow \widetilde{r}_n/|\widetilde{r}_n|\\ &\quad \textbf{end for}\\ &\textbf{end for}\end{aligned}} \tag{68}$$

where a new time index n is introduced to show that the update is not being carried out for each k but only for few (N_{i}) iterations once a batch of received data is available in memory. This new update no longer requires the phase synchronizer output z_k. Instead, the input statistic $\mathbb{E}y_k^4$ is estimated from available received batch of stored data. The normalization in the second step is required to ensure the constraint $|r_n| = 1, \forall n$. The selection of N_{i} in (68) depends upon how large is the step-size, so for a large step-size, a small N_{i} is sufficient. This steepest gradient algorithm now works without filtering to generate intermediate synchronizer output at each step. As long as the total processing time does not exceed the time it takes to collect the buffered data frame, this forward driving synchronizer will only introduce a fixed processing delay in signal processing pipeline, making it well suited not only for high data rate systems but for both fixed and variable length packet transmissions. The batch processing is highly preferable if it can support admissible performance for $N_{\text{i}} \ll N_{\text{b}}$. Likewise, we can obtain batch processing versions of those adaptive synchronizers which we introduced in Section 4.

It is interesting to note that by dropping the expectation operator in (68) and exploiting $z_k = r_k^* y_k$, we can obtain a variant of the standard FP-PRA (10) as follows:

$$\boxed{\begin{aligned}&\widetilde{r}_k = r_k - \mu\, z_k^3\, y_k\\ &r_{k+1} = \widetilde{r}_k/|\widetilde{r}_k|\end{aligned}} \tag{69}$$

The convergence performance of the *complex-valued* update (69) is a little better than that of the standard *real-valued* FP-PRA (10). Due to the lack of space, however, we leave this discussion to be addressed elsewhere.

8. Numerical simulations

8.1. Admissibility of transformation based phase recovery

In Section 4, we have shown analytically that the centroid of a uniform square moves in the direction of phase-offset under the

[12] The idea of forward driving based batch processing to realizing an iterative steepest descent algorithm is relatively new and appeared recently in the year 2012 in [59].

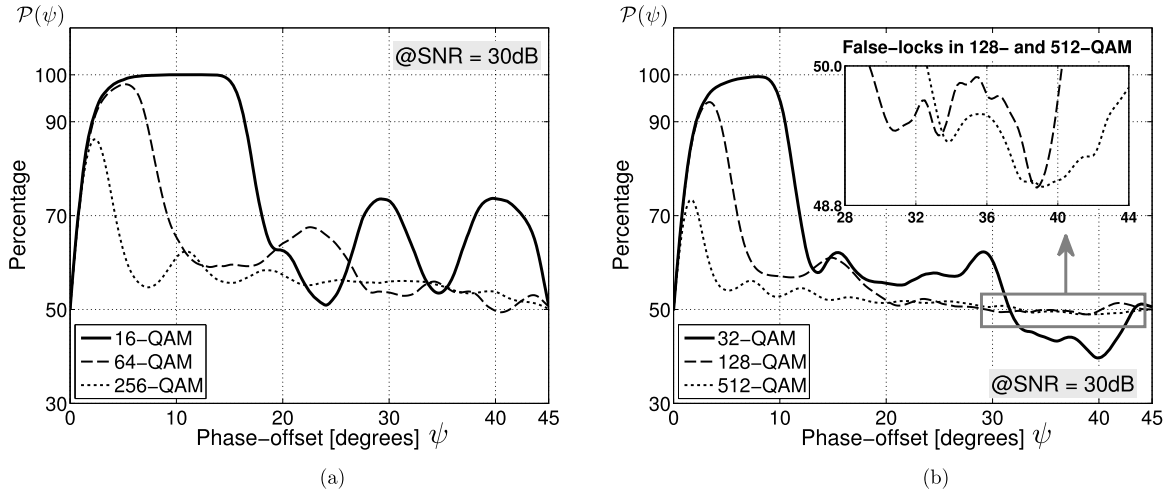

Fig. 11. The probability that the transformed symbols would maintain the true direction of movement of constellation for the phase-offset ranging from zero to 45 degrees for square-QAM. It can be noticed that all three square-QAM have this probability higher than 0.5 for all values of phase-offset.

condition $|\psi| \leq \pi/4$ radian. The movement of centroid of a uniform cross, however, creates a false lock when the phase-offset is in the range $\psi \in \{(\tan^{-1}(2/3), \pi/4] \cup (-\pi/4, -\tan^{-1}(2/3)]\}$ radian. In this numerical simulation, we aim to observe this behavior for practical square- and cross-QAM of different sizes when the QAM are transformed to QPSK (that is mapping to centroid) using the algorithm (29).

Consider the introduction of an anti-clockwise phase-offset ($0 < \psi < \pi/4$) to a noisy QAM constellation, and its transformation to QPSK (that is obtaining $\widetilde{z}_k$ from z_k for a range of values of ψ). After the transformation, due to the presence of noise, QAM alphabets do not get mapped exactly onto the ideal four-symbol QPSK constellation rather they happen to form a cluster of points around the desired QPSK symbols (or centroid in other words). Now consider the rotated symbols $\{z_k\}$ which belong to the first quadrant, $z_k \in Q_1$; denote this subset as $\{z_k(Q_1)\}$. Now consider the transformed symbols, denoted as $\{\widetilde{z}_k(Q_1)\}$, obtained from $\{z_k(Q_1)\}$; note that $\text{qam2qpsk}(z_k(Q_1)) \in Q_1$. Since we are considering a *positive* phase-offset (an anti-clockwise rotation), therefore, we are interested in those transformed symbols $\{\widetilde{z}_k(Q_1)\}$ which lie *above* the diagonal, let us denote these symbols as $\{\widetilde{z}_k(Q_1^+)\}$. Based on inspection and numerical evidences, it is found that when the cardinality of $\{\widetilde{z}_k(Q_1^+)\}$ is higher than half of the cardinality of $\{\widetilde{z}_k(Q_1)\}$ then the centroid of the cluster lies *above* the diagonal (that is in the correct direction), and the proposed phase tracker detects and corrects the phase-offset successfully. Define a ratio $\mathcal{P} \equiv \mathcal{P}(\psi)$

$$\mathcal{P} = \frac{|\{\widetilde{z}_k(Q_1^+)\}|}{|\{\widetilde{z}_k(Q_1)\}|}, \tag{70}$$

where $|\cdot|$ denotes cardinality. For a given value of ψ, if $\mathcal{P}(\psi) > 50\%$, then the algorithm is capable of detecting the true direction of phase-offset. On the other hand, the algorithm may get trapped into a false lock if $\mathcal{P}(\psi)$ happens to be less than 50%. Similarly, if $\mathcal{P}(\psi) = 50\%$, then the transformed symbols are incapable of giving any clue about the unknown phase-offset. Note that we have mentioned and described in the context of first quadrant to explain the concept, but as a matter of fact, we have evaluated the percentage (70) for all four quadrants by calculating appropriately the cardinalities $|\{\widetilde{z}_k(Q_i^+)\}|, i = 1, 2, 3, 4$. The percentage $\mathcal{P}(\psi)$ is depicted in Fig. 11 for both square- and cross-QAM which are obtained for 30 dB SNR value. Note that

1. $\mathcal{P}(0°) = \mathcal{P}(45°) = 50\%$, which is a desirable property. Not shown here (due to the lack of space) but $\mathcal{P}(-\psi) = \mathcal{P}(\psi)$ is true for all square- and cross-QAM.
2. $\mathcal{P}(\psi)$ is found to be larger than 50% for the three addressed square-QAM. Though, for 64- and 256-QAM and for $\psi > 20°$, $\mathcal{P}(\psi)$ gets close to 50% but it is not less than or equal to 50%. Also note that the range of ψ, for which the $\mathcal{P}(\psi)$ is close to 100%, gets smaller with larger QAM-sizes. Conclusively, it is admissible to apply QAM-to-QPSK transformation on the practical square-QAM receiver for the sake of carrier recovery. This is in accordance to the analysis presented in Section 4.2.1.
3. $\mathcal{P}(\psi)$ is not always larger than 50% for the three addressed cross-QAM. In the whereabout of the range 30°–44°, the plots of $\mathcal{P}(\psi)$ are below 50% which clearly indicates that centroid has moved below the diagonal and introducing a false lock. Compare these results to the analysis we carried out in the previous section in relation to the position of centroid of uniform cross – the analysis presented in Section 4.2.2 supports this observation. It appears that a QAM-to-QPSK transformation is not an admissible technique for detecting phase-offset in a cross-QAM receiver. However, owing to fluctuations due to additive noise and using a relatively larger step-size, it is possible for phase-tracker to escape such false locks as we have evidenced numerically.

In Fig. 12, we demonstrate (for the sake of further clarification that) how transformed symbols retain (on the whole) the true direction of rotation. The 16-QAM symbols lying in the first, second, third and fourth quadrants are depicted using markers ☆, ▽, ◊, and □, respectively (refer to Fig. 12(a)). In Fig. 12(b), the noisy constellation (of Fig. 12(a)) is subjected to clockwise 30 degree rotation. Fig. 12(c) depicts the result of QAM-to-QPSK transformation. Since the rotation is clockwise, we expect more symbols above the diagonal considering the first quadrant. Clearly, in the first quadrant in Fig. 12(c), some symbols (denoted by □) which entered from the fourth quadrant have all been mapped below the diagonal. Similarly, some of the ☆ symbols which happened to get clustered around the diagonal (see in Fig. 12(b)) have been mapped below the diagonal. These two set of symbols, lying below the diagonal, contribute wrong information about the rotation. However, there also exist three clusters of ☆ symbols which are lying above the diagonal, and they are larger in number than those lying below the diagonal. Due to which, the CSFP-PRA is able to detect the true direction of rotation and correct the amount of phase-offset successfully.

Now we outline a numerical simulation for cross-QAM signaling to demonstrate that it is possible for CSFP-PRA to escape false locks by using relatively larger step-sizes. When the phase synchronizer is trapped in a false lock, the MSD exhibits higher values and there is a sudden drop in its values when the phase tracker happens to escape. So we compute the probability of escape versus iterations by observing the converging values of MSD for a number of independent runs. In Fig. 11(b), we can see that a 32-QAM, when transformed to QPSK, exhibits a false lock when $36.7° \leq \psi \leq 43.6°$. So by initializing ($\phi_0 = 0$) and selecting a θ somewhere in the middle of the range $(36.7°, 43.6°)$, we can ensure a trap for synchronizer. Next, we observe convergence of escape probabilities with two step-sizes, 8×10^{-2} and 5×10^{-2}, for two different SNR values as shown in Fig. 13. Note that, at 30 dB SNR value, the synchronizer is able to escape earlier with ($\mu = 8 \times 10^{-2}$) than ($\mu = 5 \times 10^{-2}$); see cases (a) and (b). Also note that, at a lower SNR value, say 20 dB, due to the presence of higher amount of noise, the synchronizer is able to escape even earlier; see case (c).

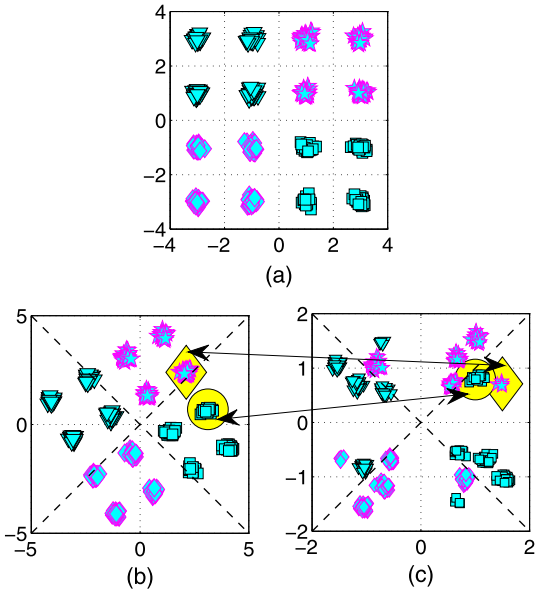

Fig. 12. Demonstrating the effect of transformation on a rotated square-QAM.

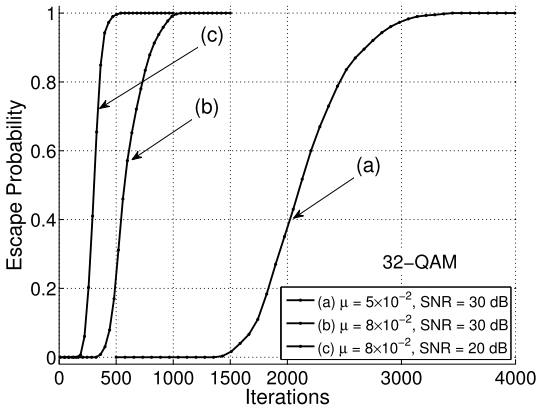

Fig. 13. Plots of escape probabilities.

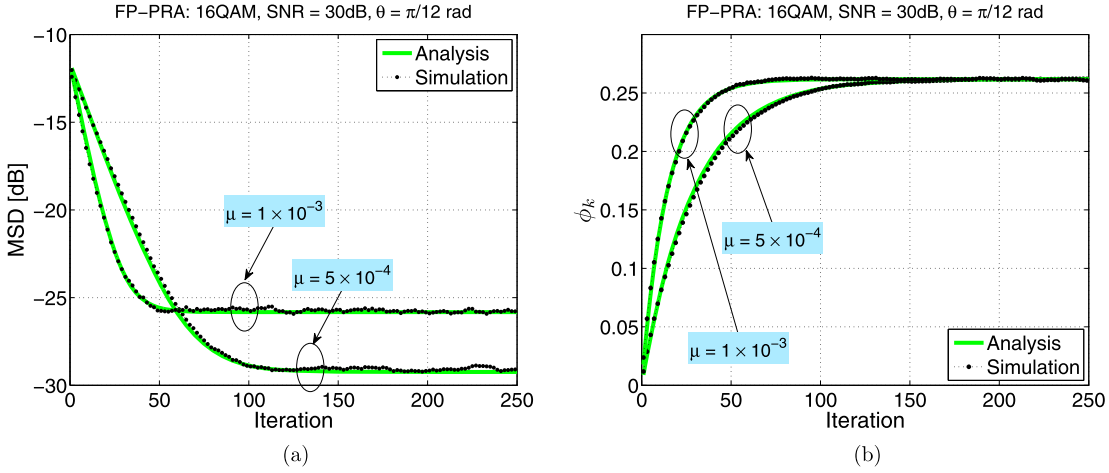

Fig. 14. FP-PRA: Analytical and simulated MSD convergence and phase acquisition traces for two different step-sizes.

8.2. Transient performance in AWGN environment

In this numerical simulation, we obtain simulated instantaneous MSD ($\mathbb{E}\psi_k^2$) and instantaneous phase estimate ($\mathbb{E}\phi_k$) and compare with their analytical counterparts for 16-QAM signaling for FP-PRA and CSFP-PRA (SGFP-PRA and SGD-PRA are skipped due to the shortage of space). In Fig. 14 and 15, we show results obtained for FP-PRA and CSFP-PRA, respectively, with 30 dB SNR value in a stationary environment ($\sigma_q = 0$) with no frequency mismatch ($\Omega = 0$). It can clearly be seen that the simulated and analytical traces exhibit good conformation. Also notice that, for the same number of iterations required for the convergence, CSFP-PRA performed far better than the standard FP-PRA.

8.3. Steady-state performance in AWGN environment

In this numerical simulation, we obtain simulated steady-state MSD and compare with their analytical counterparts with both square- and cross-QAM signaling for FP-PRA, SGD-PRA, SGFP-PRA, and CSFP-PRA. In Fig. 16, we show results obtained for 16-, and 64-QAM for different SNR values in stationary environment ($\sigma_q = 0$) with no frequency mismatch ($\Omega = 0$).

1. The step-size for CSFP-PRA is selected to be 8×10^{-3} for all QAM sizes. The step-sizes of other two algorithms, FP-PRA and SGFP-PRA, have been selected such that all three algorithms converge requiring similar number of iterations. The convergence time (at SNR $= 30$ dB) is noted to be nearly 380, and 800 iterations for 16-, and 64-QAM, respectively.
2. For the evaluation of statistical parameters c_1, c_2 and c_3 (for the given values of SNR and step-size), we consider all QAM alphabets for FP-PRA, selected alphabets (belonging to go-region) for SGFP-PRA, and four alphabets $\pm 1 \pm 1\mathbf{i}$ for CSFP-PRA.

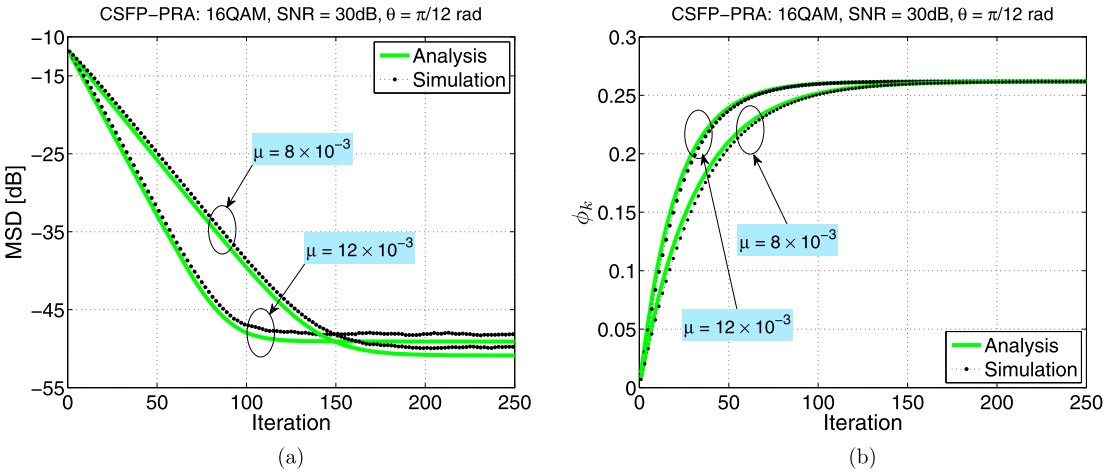

Fig. 15. CSFP-PRA: Analytical and simulated MSD convergence and phase acquisition traces for two different step-sizes.

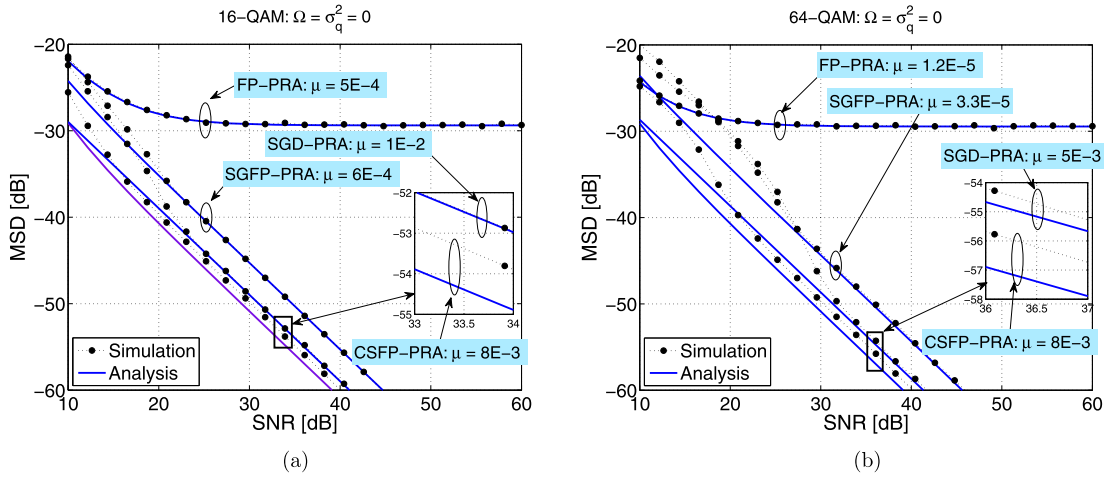

Fig. 16. MSD versus SNR for 16- and 64-QAM.

3. The performance of FP-PRA is not improved at higher values of SNR and it attempts to attain an MSD floor. On the other hand, SGFP-PRA and CSFP-PRA, are demonstrating improved performance with an increase in SNR, which is a desirable property.
4. Note that CSFP-PRA is performing better than the standard FP-PRA for all values of SNR (i.e., 10 dB and above). The SGFP-PRA, however, exhibits relatively inferior performance (to CSFP-PRA) at higher QAM sizes due to the availability of very few go-regions to detect phase-offset and update the tracking algorithm.
5. The analytical MSD are very close to numerical ones for all SNR values for FP-PRA. For SGFP-PRA and CSFP-PRA, however, analytical and numerical MSD are close only when SNR values are greater than 20 dB.

Similarly, in Fig. 17, we provide plots of MSD versus SNR for (cross) 32- and 128-QAM for the four addressed algorithms. We first notice that the simulated and analytical results are close only when SNR values are at higher side (say above 20 dB) for 32-QAM. For 128-QAM, on the other hand, the simulated and analytical MSD plots are close only for FP-PRA. So unlike square-QAM, where we have seen a close agreement between analytical findings and simulation results, the analytical and the simulated MSD values for cross-QAM are far apart from each other. A major reason behind this behavior is that, unlike the square-QAM, the cross-QAM do not have *independent* in-phase and quadrature components.[13] Further note that CSFP-PRA outperforms FP-PRA when SNR is greater than 20 dB for both 32- and 128-QAM. For 128-QAM with SNR value than 20 dB, the performance of SGFP-PRA and SGD-PRA deteriorates very much as compared to FP-PRA. A strong reason for the failure of SGFP-PRA and SGD-PRA is the poor availability of 'go' zones for reliable updates at lower SNR values.

8.4. Performance evaluation with FO and/or PN

In this set of numerical simulations, we first evaluate the performances of FP-PRA and CSFP-PRA in the presence of FO for 16- and 64-QAM at 30 dB SNR value. So we obtain the values of MSD (analytically as well as numerically) by changing Ω from 1×10^{-5} to 1×10^{-1}. Results are depicted in Fig. 18. Both algorithms are run using optimal values of step-size as specified in (41). Note that both theory and simulation are in good conformation. Secondly the

[13] The problem of blind phase recovery has close affinity with independent component analysis (ICA) [60]. From the theory of ICA, we understand that when complex-valued sources have dependent or correlated in-phase and quadrature components, then the fourth-order statistics in the cost function must also include correlation- and improperness-sensitive terms; see [60, Section 6.10.7.3] and references therein. Since, a cross-QAM has dependent quadrature components, therefore, the ICA methods for dependent complex-valued sources, as in [61–64], may help in designing improved phase recovery algorithms for cross-QAM. We leave this discussion to be explored elsewhere.

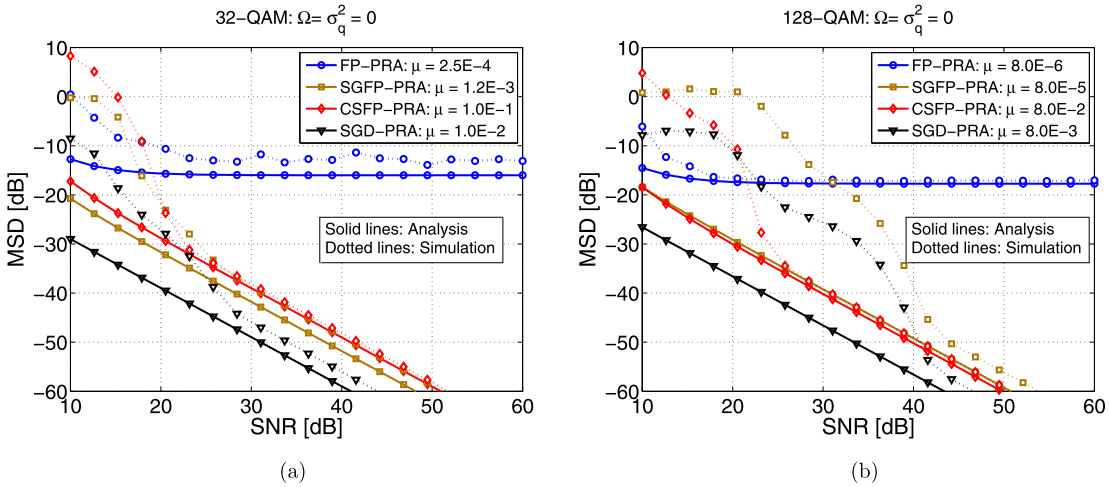

Fig. 17. MSD versus SNR for 32- and 128-QAM.

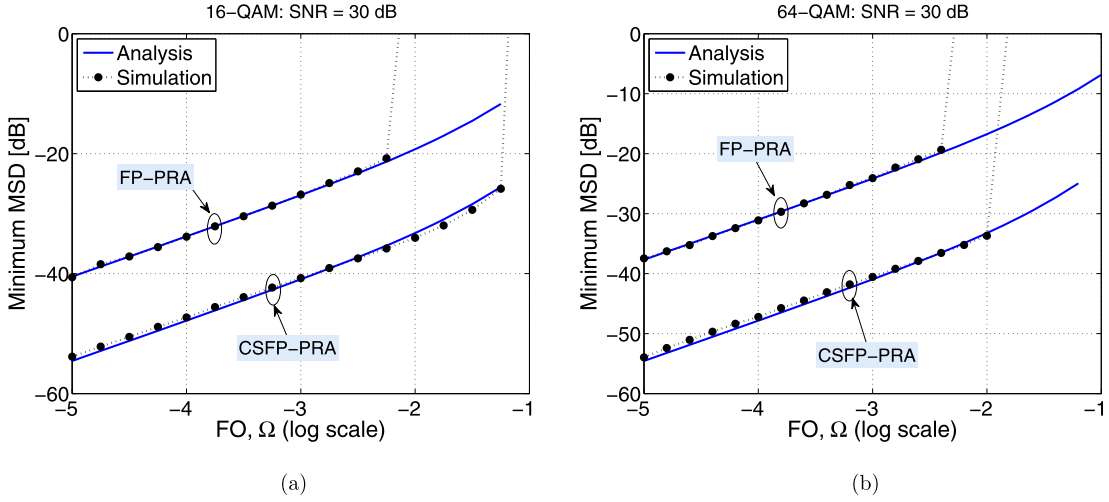

Fig. 18. Minimum MSD versus FO for 16- and 64-QAM obtained with the optimal step-sizes.

MSD of CSFP-PRA is lower by nearly 13 dB as compared to that of FP-PRA. Thirdly CSFP-PRA can be seen to be able to tolerate much higher FO. Like for 16-QAM, the maximum FO which is tolerated by FP-PRA and CSFP-PRA (beyond which they diverged) is 0.0056 and 0.0562, respectively. Similarly, for 64-QAM, the maximum FO which is tolerated by FP-PRA and CSFP-PRA is 0.004 and 0.01, respectively.

In the second set of numerical simulations, we evaluate the performances of FP-PRA and CSFP-PRA in the presence of PN for 16- and 64-QAM at 30 dB SNR value. We obtain the values of MSD (analytically as well as numerically) by changing σ_q^2 from 1×10^{-5} to 1×10^{-2}. Results are depicted in Fig. 19. Both algorithms are run using optimal values of step-size as specified in (46). Note that both theory and simulation are in good conformation.[14] Secondly the MSD of CSFP-PRA is lower by nearly 10 dB as compared to that of FP-PRA. Thirdly CSFP-PRA can be seen to be able to track PN with much larger variance like, for 16-QAM, the maximum σ_q^2 which can be tracked by FP-PRA and CSFP-PRA (beyond which they diverged) is 1.77×10^{-4} and 3.20×10^{-3}, respectively. For 64-QAM, the maximum σ_q^2 which can be tracked by FP-PRA and CSFP-PRA is close to 7.94×10^{-4}.

In the next numerical simulation, we demonstrate that, in the presence of FO and/or PN, the MSD is a convex downward function of the step-size (loop gain) and there exists an optimal value of step-size for which MSD is minimized, which depending on the scenario, can be obtained from (41), (46), or (49). We consider signaling of 16- and 64-QAM at 30 dB SNR each under two scenarios of PN and FO using synchronizers FP-PRA and CSFP-PRA. The results are depicted in Fig. 20, where it can been that our analytical findings are in good conformation with simulation results. Note that each simulation point is obtained as an average of 1000 independent runs with random initialization of phase-offset, noise and signal points. The steady-state MSD is evaluated, in each run, by averaging estimated MSD values for 100,000 iterations once the convergence is achieved.

At this stage, we would like to quote from [38]:

If there is a residual frequency-offset or large amount of phase noise, the constellation would be misaligned with the transformation boundaries. Some of the symbols would then be transformed in the wrong direction and result in an incorrect phase noise estimate. Therefore, this QPSK transformation algorithm can only be used after compensation for the residual frequency-offset between the transmit-

[14] For QPSK, we have found that the simulated and the analytical MSD values of FP-PRA differ significantly for higher values of phase noise variance (PNV); results of which are omitted here due to the lack of space. Since CSFP-PRA transforms any given QAM constellation into QPSK, therefore, its simulated and analytical MSD values happen to differ from each other at higher PNV as shown in Fig. 19.

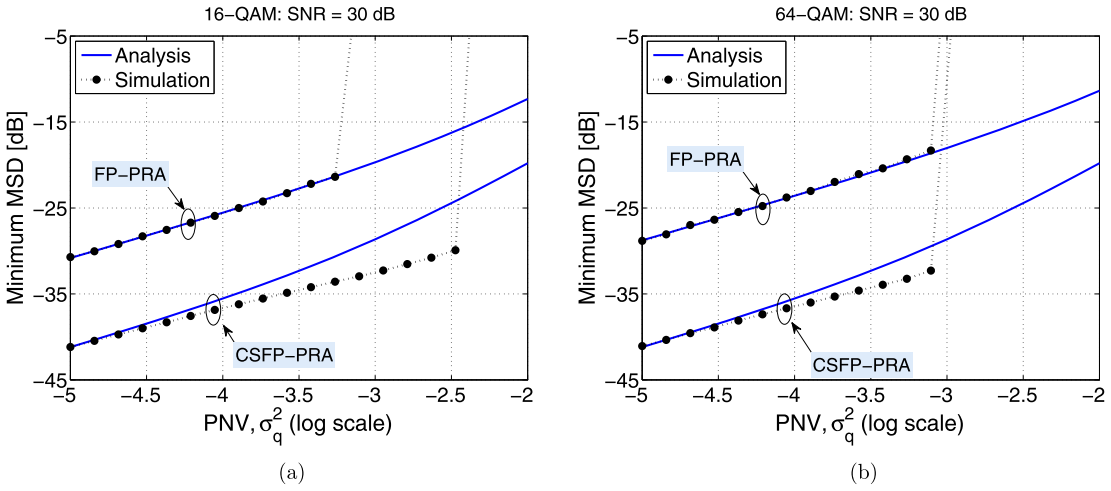

Fig. 19. Minimum MSD versus PN variance (PNV) for 16- and 64-QAM obtained with the optimal step-sizes.

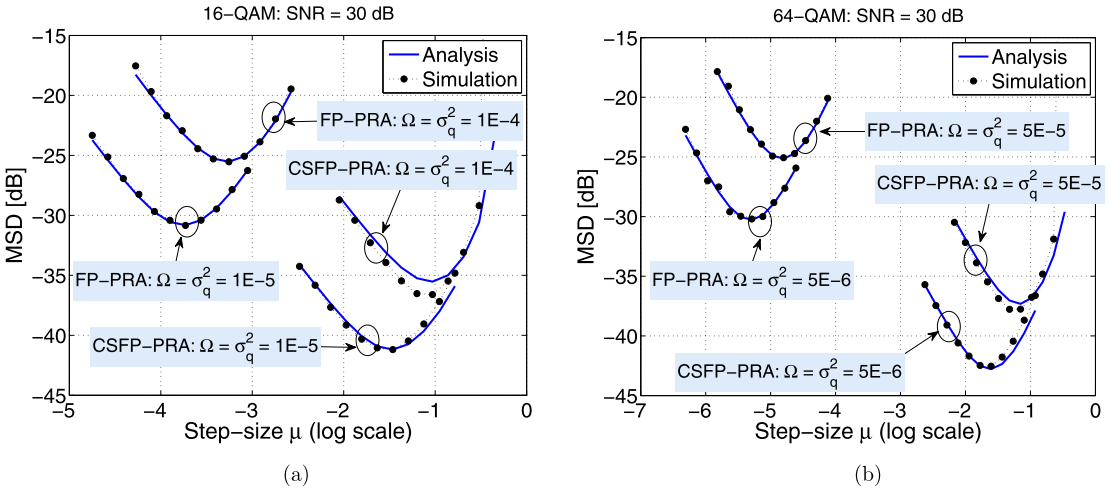

Fig. 20. Minimum MSD versus step-size for 16- and 64-QAM subjected to FO and PN.

ter and local oscillator lasers and an initial correction for the laser phase noise using a coarse estimate.

In contrast, we have seen that our addressed adaptive synchronizers (FP-PRA and CSFP-PRA) have acquired successfully the phase-offset in the presence of frequency-offset and phase noise.

8.5. Performance evaluation with ISI

In this numerical simulation, we evaluate the performances of FP-PRA for 16-QAM in the presence of ISI (but no noise), and $\theta = \pi/300$ rad. We assume that the joint channel-equalizer impulse response contains two taps: $h_\delta = 1$, and $h_{\delta+1} = 0.21 + \mathbf{i}0.12$, (where δ is arbitrary), and for which we have $\omega = 7.5 \times 10^{-4}$ rad. Results are depicted in Fig. 21 which shows that ϕ_k converges to $\theta + \omega$ thus incorporating the effects of both fixed rotation θ as well as the rotation caused by the interfering path. The figure also illustrates the constellation of the received signal in the presence of ISI. Note that the traces are obtained as an ensemble average of 8000 independent runs.

8.6. Adaptive estimation of optimal step-size

In this numerical simulation, we demonstrate the performances of VSSFP-PRA and VSSCSFP-PRA for 16-QAM at SNR equal to 30 dB

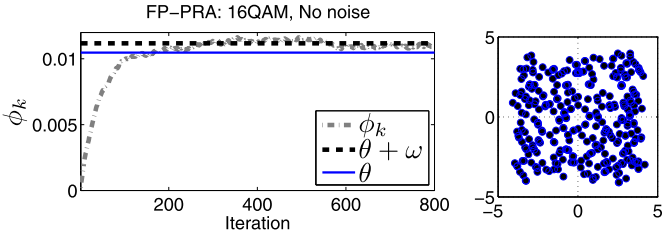

Fig. 21. Effect of ISI on the steady-state convergence of FP-PRA (left). ISI distorted constellation (right).

Table 1
Parameters for VSS simulation.

	VSSFP-PRA	VSSCSFP-PRA
α	1/40	1/10
$\mu_{\min}$	2×10^{-6}	3×10^{-4}
$\mu_{\max}$	2×10^{-3}	1

in the presence of phase noise. We assume that the phase noise follows the model (43) with $\sigma_q^2 = 1 \times 10^{-5}$ for which the optimal step-sizes, realizing FP-PRA and CSFP-PRA, are given as 1.78×10^{-4} and 3.29×10^{-2}, respectively. Both algorithms, VSSFP-PRA and VSSCSFP-PRA, are initialized as $\mu_0 = \varpi_0 = 0$. The values of α, $\mu_{\min}$ and $\mu_{\max}$ are summarized in Table 1. To assess the performance,

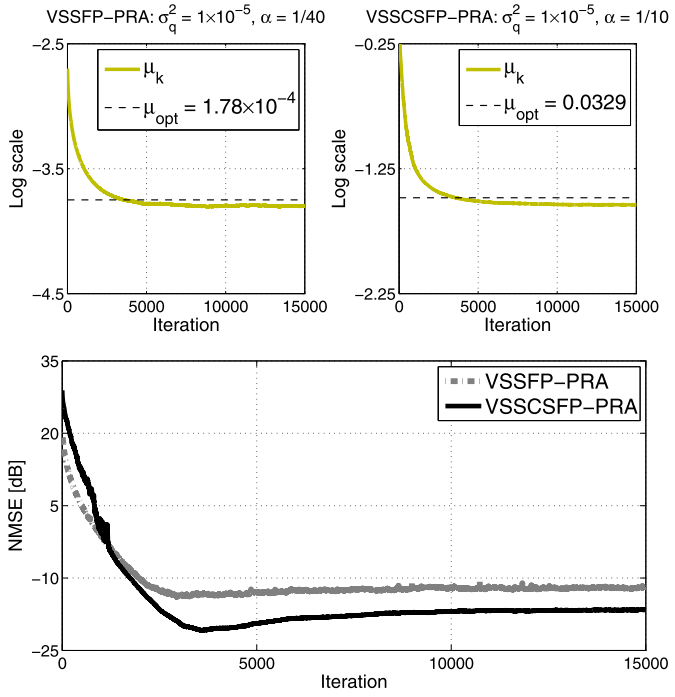

Fig. 22. Results for VSS simulation.

we use the following normalized mean squared error (NMSE) metric for the estimated μ_k:

$$\text{NMSE}_k = \frac{1}{N_{\text{run}}} \sum_{j=1}^{N_{\text{run}}} \left(\frac{\mu_k(j) - \mu_{\text{opt}}}{\mu_{\text{opt}}} \right)^2, \tag{71}$$

where NMSE_k is the value of NMSE at the kth iteration which is obtained as an ensemble average of N_{run} independent runs, and $\mu_k(j)$ is the value of variable μ in kth iteration of the jth run. Here, we consider $N_{\text{run}} = 1000$, and obtain the traces of μ_k and NMSE_k for both VSSFP-PRA and VSSCSFP-PRA as depicted in Fig. 22. Clearly, both algorithms can be noticed to converge successfully in the vicinity of their optimal values, where the steady-state NMSE floor of VSSCSFP-PRA is lower than that of VSSFP-PRA.

In this numerical simulation, VSSFP-PRA is found to have converged to 1.6×10^{-4} with standard deviation $\pm 1.89 \times 10^{-5}$ while the true optimum value is 1.78×10^{-4}. Similarly, VSSCSFP-PRA converges to 0.0289 with standard deviation ± 0.0016 while the true optimum is 0.0329. These results show that the steady-state values of adaptive step-sizes are slightly smaller than the true values. At this stage, however, it is not completely clear to us whether the sparse optimization is responsible for this bias. A possible reason, which we have figured out, is stuck-at-$\mu_{\min}$ problem, i.e., both VSSFP-PRA and VSSCSFP-PRA have the tendency to get stuck at $\mu_{\min}$ with a non-zero probability. Like, in the above numerical simulation, in one or two out of 1000 independent trials, the algorithms have been found to get stuck at $\mu_{\min}$. Possible remedies to this problem include re-initialization, not to select very small $\mu_{\min}$, or use *ordinary* gradient method at the start-up and move gradually towards *natural* gradient (sparse) method during the course of convergence. Such a *convex* combination is possible and is expected to resolve the matter. Convex combinations of two adaptive filters are well understood [65] and have been employed recently in the similar (i.e., sparse filtering) context in [66].

8.7. Performance of batch FP-PRA

In this section, the performances of standard FP-PRA and its batch version (67)–(68) are presented. Simulated MSD traces for

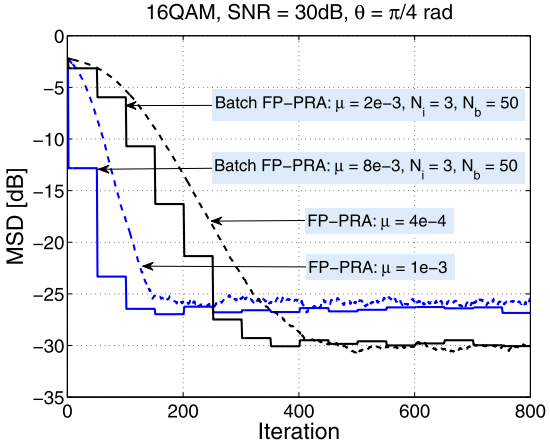

Fig. 23. Comparison between the standard and the batch FP-PRA.

16-QAM signaling at 30 dB SNR value with $\theta = \pi/4$ by Monte-Carlo method using 1000 independent runs are obtained. The obtained results are shown in Fig. 23, where batch FP-PRA can be seen to be performing slightly better than FP-PRA when allowed to reach same steady-state MSD floor. Most importantly, at each stage, i.e., after every N_{b} number of received symbols, the batch FP-PRA (equipped with relatively larger step-size) required only three iterations (i.e., $N_{\text{i}} = 3$) in a forward driving manner while using the statistics obtained from last fifty received symbols (i.e., $N_{\text{b}} = 50$). Noticeably, the batch processing can be seen to have the potential to offer considerable saving in computations.

9. Conclusions

In this work, we revisited the fourth-power adaptive phase recovery (FP-PRA) algorithm to develop three new variants for improved performance in situations suffering with phase mismatch, frequency mismatch, and phase noise for QAM based communication systems. We analyzed the dynamic, steady-state and tracking performances of all four addressed algorithms, and obtained the optimal values of step-size (loop gain) for the given scenarios. The effect of imperfect equalization on phase synchronization was also studied. Simulation results were found to be in good conformation with the analytical findings.

Specifically, one of the proposed algorithms, CSFP-PRA, was based on the transformation of QAM symbols so that they are mapped onto a QPSK constellation. We discussed this existing transformation idea with a new interpretation which relates the process of mapping to finding the quadrant-wise *centroid* of the QAM constellation, and how the movement of centroid helped in determining the unknown rotation. Based on our analysis, we have been able to show that QAM-to-QPSK transformation is an admissible strategy for finding the unknown phase rotation in square-QAM. In the case of cross-QAM, however, we have discovered existence of false locks. Using numerical simulations, we demonstrated how a larger step-size can help escape such false locks (in cross-QAM). A generic iterative method for the transformation of given QAM into QPSK was also presented.

The other two proposed algorithms, SGD-PRA and SGFP-PRA, used different sensitivity functions in their updates but shared a similar *stop-and-go* strategy. Such update was allowed only when rotated QAM symbols fall in specific annular regions; otherwise, the update process was stopped. These annular regions in (rotation free) QAM constellation were selected such that they contained QPSK-like symbols. For successful operation of these algorithms, however, it was required to ensure the perfect recovery of true energy of the signal (which can be facilitated easily by DSP

based automatic gain control). The relative performance of CSFP-PRA has been found to be better than the other three algorithms, i.e., FP-PRA, SGFP-PRA, and SGD-PRA, for all sizes of square-QAM and large cross-QAM. For 32-QAM, the SGD-PRA is found to be outperforming the others. For 256- and higher square/cross-QAM, the performance of SGFP-PRA, and SGD-PRA deteriorated significantly due to very small reliable zones available for the update (results for 256- and higher QAM have been skipped due to the lack of space). Moreover, detailed numerical simulations revealed that CSFP-PRA can tolerate larger frequency-offset and phase noise variance (with far lower MSD) than FP-PRA.

We also presented an adaptive (variable) procedure to obtain optimal values of step-size for the standard FP-PRA and CSFP-PRA, which resulted in two variable step-size algorithms VSSFP-PRA and VSSCSFP-PRA for systems exhibiting phase noise. The Kivinen–Warmuth method for sparse optimization has been exploited to obtain approximate natural gradient update for variable step-sizes. Simulation results showed good convergence in the vicinity of optimal values without requiring statistical knowledge of phase noise. The second-order realization of these algorithms for the compensation of frequency-offset was considered straightforward and has been left for others to explore.

The last part of this work presented a batch processing realization of adaptive phase recovery. It exploited the idea of forward driving and modified the existing stochastic gradient real-valued update into an iterative steepest descent type complex-valued update. Simulation results showed that the proposed batch processing can offer a hardware efficient solution for synchronization in higher data rate systems.

Acknowledgments

We acknowledge the support of COMSATS Institute of Information Technology (Islamabad campus), Pakistan, King Fahd University of Petroleum and Minerals, Saudi Arabia, and Brunel University London, UK, towards the accomplishment of this work. We also acknowledge the Editor and anonymous reviewers for their valuable feedback.

Appendix A. Evaluation of centroid for uniform square

Consider a clockwise rotation by an angle $\psi < \pi/4$, note that

$$\tan\left(\chi + \frac{\pi}{4}\right) = \frac{b_C}{a_C} \;\Rightarrow\; \chi = \tan^{-1}\left(\frac{b_C - a_C}{b_C + a_C}\right). \tag{72}$$

For a uniform square, we obtain:

$$\begin{aligned} a_C &= \int_0^{\cos(\psi)-\sin(\psi)} xS_{x,1}\mathrm{d}x + \int_{\cos(\psi)-\sin(\psi)}^{\sec(\psi)} xS_{x,2}\mathrm{d}x \\ &= \frac{2\left(\cos(\psi)^3 + \sin(\psi)^3\right) + \cos(\psi) - \sin(\psi)}{6\cos(\psi)^2} \end{aligned} \tag{73}$$

where the boundaries $S_{x,1}$ and $S_{x,2}$, as illustrated in Fig. 24(a), are expressed as:

$$S_{x,1} = \frac{\cos(\psi) + \sin(\psi) - \sec(\psi)}{\cos(\psi) - \sin(\psi)} x + \sec(\psi), \tag{74a}$$

$$S_{x,2} = \frac{(\cos(\psi) + \sin(\psi))(x - \sec(\psi))}{\cos(\psi) - \sin(\psi) - \sec(\psi)}. \tag{74b}$$

Next b_C is obtained

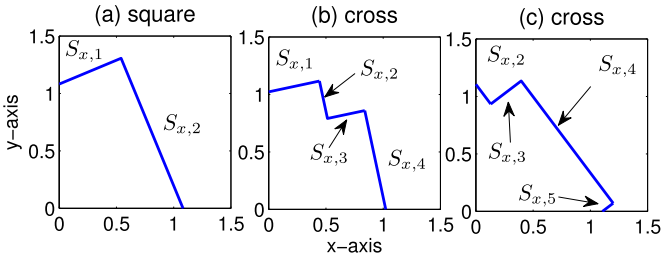

Fig. 24. Geometries of rotated square and cross.

$$\begin{aligned} b_C &= \tfrac{1}{2}\int_0^{\cos(\psi)-\sin(\psi)} S_{x,1}^2\mathrm{d}x + \tfrac{1}{2}\int_{\cos(\psi)-\sin(\psi)}^{\sec(\psi)} S_{x,2}^2\mathrm{d}x, \\ &= \frac{2\left(\cos(\psi)^3 - \sin(\psi)^3\right) + \cos(\psi) + \sin(\psi)}{6\cos(\psi)^2}. \end{aligned} \tag{75}$$

Combining the above findings, we obtain a closed form relation between phase error ψ and angle of centroid χ (as measured from diagonal) as specified in (24).

Appendix B. Evaluation of centroid for uniform cross

Before false-lock, we have for a uniform cross

$$\begin{aligned} a_C &= \tfrac{9}{8}\int_0^{\frac{2}{3}\cos(\psi)-\sin(\psi)} xS_{x,1}\mathrm{d}x + \tfrac{9}{8}\int_{\frac{2}{3}\cos(\psi)-\sin(\psi)}^{\frac{2}{3}(\cos(\psi)-\sin(\psi))} xS_{x,2}\mathrm{d}x \\ &\quad + \tfrac{9}{8}\int_{\frac{2}{3}(\cos(\psi)-\sin(\psi))}^{\cos(\psi)-\frac{2}{3}\sin(\psi)} xS_{x,3}\mathrm{d}x + \tfrac{9}{8}\int_{\cos(\psi)-\frac{2}{3}\sin(\psi)}^{\sec(\psi)} xS_{x,4}\mathrm{d}x \\ &= \frac{13\cos(\psi)^3 + 9\cos(\psi) + 13\sin(\psi)^3 - 4\sin(\psi)}{48\cos(\psi)^2} \end{aligned} \tag{76}$$

where the boundaries $S_{x,1}$, $S_{x,2}$, $S_{x,3}$ and $S_{x,4}$, as illustrated in Fig. 24(b), are expressed as:

$$S_{x,1} = \frac{\cos(\psi) + \frac{2}{3}\sin(\psi) - \sec(\psi)}{\frac{2}{3}\cos(\psi) - \sin(\psi)} x + \sec(\psi), \tag{77a}$$

$$\begin{aligned} S_{x,2} &= \cot(\psi)\left(\tfrac{2}{3}\cos(\psi) - \sin(\psi) - x\right) \\ &\quad + \cos(\psi) + \tfrac{2}{3}\sin(\psi), \end{aligned} \tag{77b}$$

$$\begin{aligned} S_{x,3} &= \tan(\psi)\left(x - \tfrac{2}{3}(\cos(\psi) - \sin(\psi))\right) \\ &\quad + \tfrac{2}{3}(\cos(\psi) + \sin(\psi)), \end{aligned} \tag{77c}$$

$$S_{x,4} = \frac{\left(\frac{2}{3}\cos(\psi) + \sin(\psi)\right)(x - \sec(\psi))}{\cos(\psi) - \frac{2}{3}\sin(\psi) - \sec(\psi)}. \tag{77d}$$

Next we obtain b_C as follows:

$$\begin{aligned} b_C &= \tfrac{9}{16}\int_0^{\frac{2}{3}\cos(\psi)-\sin(\psi)} S_{x,1}^2\mathrm{d}x + \tfrac{9}{16}\int_{\frac{2}{3}\cos(\psi)-\sin(\psi)}^{\frac{2}{3}(\cos(\psi)-\sin(\psi))} S_{x,2}^2\mathrm{d}x \\ &\quad + \tfrac{9}{16}\int_{\frac{2}{3}(\cos(\psi)-\sin(\psi))}^{\cos(\psi)-\frac{2}{3}\sin(\psi)} S_{x,3}^2\mathrm{d}x + \tfrac{9}{16}\int_{\cos(\psi)-\frac{2}{3}\sin(\psi)}^{\sec(\psi)} S_{x,4}^2\mathrm{d}x \\ &= \frac{13\cos(\psi)^3 + 9\cos(\psi) - 13\sin(\psi)^3 + 4\sin(\psi)}{48\cos(\psi)^2} \end{aligned} \tag{78}$$

During the false-lock, the geometry follows Fig. 24(c); we obtain

$$a_C = \tfrac{9}{8}\int_0^{\frac{2}{3}(\cos(\psi)-\sin(\psi))} xS_{x,2}\,dx + \tfrac{9}{8}\int_{\frac{2}{3}(\cos(\psi)-\sin(\psi))}^{\cos(\psi)-\frac{2}{3}\sin(\psi)} xS_{x,3}\,dx + \tfrac{9}{8}\int_{\cos(\psi)-\frac{2}{3}\sin(\psi)}^{\cos(\psi)+\frac{2}{3}\sin(\psi)} xS_{x,4}\,dx - \tfrac{9}{8}\int_{\frac{2}{3}(\cos(\psi)-\sin(\psi))\cot(\psi)+\frac{2}{3}(\cos(\psi)+\sin(\psi))}^{\cos(\psi)+\frac{2}{3}\sin(\psi)} xS_{x,5}\,dx = \frac{-19\cos(\psi)^3+17\cos(\psi)+4\sin(\psi)^3+2\sin(\psi)}{36\sin(\psi)^2} \quad (79)$$

where $S_{x,2}$, $S_{x,3}$ and $S_{x,4}$ are as specified in (77a)–(77d), while $S_{x,5}$ is defined as follows:

$$S_{x,5} = \tan(\psi)\left(x - \tfrac{2}{3}(\cos(\psi)+\sin(\psi))\right) + \tfrac{2}{3}(\sin(\psi)-\cos(\psi)). \quad (80)$$

Next we obtain b_C as follows:

$$b_C = \tfrac{9}{16}\int_0^{\frac{2}{3}(\cos(\psi)-\sin(\psi))} S_{x,2}^2\,dx + \tfrac{9}{16}\int_{\frac{2}{3}(\cos(\psi)-\sin(\psi))}^{\cos(\psi)-\frac{2}{3}\sin(\psi)} S_{x,3}^2\,dx + \tfrac{9}{16}\int_{\cos(\psi)-\frac{2}{3}\sin(\psi)}^{\cos(\psi)+\frac{2}{3}\sin(\psi)} S_{x,4}^2\,dx - \tfrac{9}{16}\int_{\frac{2}{3}(\cos(\psi)-\sin(\psi))\cot(\psi)+\frac{2}{3}(\cos(\psi)+\sin(\psi))}^{\cos(\psi)+\frac{2}{3}\sin(\psi)} S_{x,5}^2\,dx = \frac{4\cos(\psi)^3-2\cos(\psi)+19\sin(\psi)^3+2\sin(\psi)}{36\sin(\psi)^2}. \quad (81)$$

Appendix C. MSD of FP-PRA

Equating odd-order moments to zero, we obtain $\mathbb{E}[\mathcal{S}_k^2|\psi_k] = \mathbb{E}\big[x_I^4(e_I^4\cos(2\psi) - \tfrac{17}{8}e_I^4\cos(4\psi) - e_R^4\cos(2\psi) - \tfrac{17}{8}e_R^4\cos(4\psi) - \tfrac{35}{16}x_R^4\cos(8\psi) + \tfrac{9}{8}e_I^4 + \tfrac{9}{8}e_R^4 + \tfrac{3}{16}x_R^4 + \tfrac{9}{4}e_I^2e_R^2 + \tfrac{3}{2}e_I^2x_R^2 + \tfrac{3}{2}e_R^2x_R^2 + \tfrac{51}{4}e_I^2e_R^2\cos(4\psi) + \tfrac{3}{8}e_I^2x_R^2\cos(2\psi) + \tfrac{105}{8}e_I^2x_R^2\cos(6\psi) - \tfrac{3}{8}e_R^2x_R^2 \times \cos(2\psi) - \tfrac{105}{8}e_R^2x_R^2\cos(6\psi)) - \tfrac{1}{32}x_R^8\cos(8\psi) - x_I^8(\tfrac{1}{32}\cos(8\psi) - \tfrac{1}{32}) + \tfrac{1}{32}x_R^8 + x_I^6(\tfrac{3}{8}e_I^2\cos(2\psi) - \tfrac{7}{8}e_I^2\cos(6\psi) - \tfrac{3}{8}e_R^2\cos(2\psi) + \tfrac{7}{8}e_R^2 \times \cos(6\psi) + \tfrac{7}{8}x_R^2\cos(8\psi) + \tfrac{1}{2}e_I^2 + \tfrac{1}{2}e_R^2 + \tfrac{1}{8}x_R^2) + e_I^2e_R^6 - 2e_I^4e_R^4 + e_I^6e_R^2 + \tfrac{1}{2}e_I^2x_R^6 + \tfrac{9}{8}e_I^4x_R^4 + \tfrac{1}{2}e_I^6x_R^2 + \tfrac{1}{2}e_R^2x_R^6 + \tfrac{9}{8}e_R^4x_R^4 + \tfrac{1}{2}e_R^6x_R^2 + x_I^2(\tfrac{1}{2}e_R^6\cos(2\psi) - \tfrac{1}{2}e_I^6\cos(2\psi) + \tfrac{7}{8}x_R^6\cos(8\psi) + \tfrac{1}{2}e_I^6 + \tfrac{1}{2}e_R^6 + \tfrac{1}{8}x_R^6 + \tfrac{3}{2}e_I^2e_R^4 + \tfrac{3}{2}e_I^4e_R^2 + \tfrac{3}{2}e_I^2x_R^4 + \tfrac{9}{4}e_I^4x_R^2 + \tfrac{3}{2}e_R^2x_R^4 + \tfrac{9}{4}e_R^4x_R^2 - \tfrac{27}{2}e_I^2e_R^4 \times \cos(2\psi) + \tfrac{27}{2}e_I^4e_R^2\cos(2\psi) - \tfrac{3}{8}e_I^2x_R^4\cos(2\psi) + \tfrac{51}{4}e_I^4x_R^2\cos(4\psi) - \tfrac{105}{8}e_I^2x_R^4\cos(6\psi) + \tfrac{3}{8}e_R^2x_R^4\cos(2\psi) + \tfrac{51}{4}e_R^4x_R^2\cos(4\psi) + \tfrac{105}{8}e_R^2x_R^4 \times \cos(6\psi) + \tfrac{9}{2}e_I^2e_R^2x_R^2 - \tfrac{153}{2}e_I^2e_R^2x_R^2\cos(4\psi)) - \tfrac{3}{8}e_I^2x_R^6\cos(2\psi) - e_I^4x_R^4 \times \cos(2\psi) + \tfrac{1}{2}e_I^6x_R^2\cos(2\psi) - \tfrac{17}{8}e_I^4x_R^4\cos(4\psi) + \tfrac{7}{8}e_I^2x_R^6\cos(6\psi) + \tfrac{3}{8}e_R^2x_R^6\cos(2\psi) + e_R^4x_R^4\cos(2\psi) - \tfrac{1}{2}e_R^6x_R^2\cos(2\psi) - \tfrac{17}{8}e_R^4x_R^4\cos(4\psi) - \tfrac{7}{8}e_R^2x_R^6\cos(6\psi) + \tfrac{9}{4}e_I^2e_R^2x_R^4 + \tfrac{3}{2}e_I^2e_R^4x_R^2 + \tfrac{3}{2}e_I^4e_R^2x_R^2 + \tfrac{27}{2}e_I^2e_R^4x_R^2 \times \cos(2\psi) - \tfrac{27}{2}e_I^4e_R^2x_R^2\cos(2\psi) + \tfrac{51}{4}e_I^2e_R^2x_R^4\cos(4\psi)\big]$.

After some algebra, we obtain

$$\mathbb{E}[\mathcal{S}_k^2|\psi_k] = (2\mathbb{E}x_R^8 + 70\mathbb{E}x_I^4x_R^4 - 56\mathbb{E}x_I^6x_R^2)\psi_k^2 + 12\sigma_e^8 + 2\mathbb{E}x_I^2x_R^6 - 2\mathbb{E}x_I^4x_R^4 + 18\sigma_e^4\mathbb{E}x_R^4 + 2\sigma_e^2\mathbb{E}x_R^6 + 48\sigma_e^6\mathbb{E}x_R^2 + 6\sigma_e^2\mathbb{E}x_I^2x_R^4 + 18\sigma_e^4\mathbb{E}x_I^2x_R^2, \quad (82)$$

where we exploit Taylor's series simplifications $\sin(\psi)\approx\psi$ and $\cos(\psi)\approx 1-0.5\psi^2$, and statistics $\mathbb{E}e_R^2 = \sigma_e^2$, $\mathbb{E}e_R^4 = 3\sigma_e^4$, and $\mathbb{E}e_R^6 = 15\sigma_e^6$ for the quadrature components of additive Gaussian noise. Similarly, assuming $\sin(\psi)\approx\psi$, and $\cos(\psi)\approx 1$, we obtain

$$\mathbb{E}[\mathcal{S}_k|\psi_k] = \psi_k\mathbb{E}[6x_I^2x_R^2 - x_R^4 - x_I^4]. \quad (83)$$

Substituting the values of moments $\mathbb{E}[\mathcal{S}_k^2|\psi_k]$ and $\mathbb{E}[\mathcal{S}_k|\psi_k]$ in (32), we obtain the MSD expression for FP-PRA.

Appendix D. MSD of SGD-PRA

Here we evaluate MSD expression for the algorithm (22) whose sensitivity function is given by

$$\mathcal{S}_k = \Im\big[z_k\,\mathrm{csign}(z_k^*)\big] = z_{k,I}\mathrm{sign}[z_{k,R}] - z_{k,R}\mathrm{sign}[z_{k,I}] \quad (84)$$

Owing to variance relation (32), the MSD may be obtained by solving the expression: $\lim_{k\to\infty}\mathbb{E}\mathcal{S}_k^2 = \frac{2}{\mu}\mathbb{E}\psi_k\mathcal{S}_k$. We start with $\mathbb{E}\psi_k\mathcal{S}_k = \mathbb{E}[\psi_k\mathbb{E}[\mathcal{S}_k|\psi_k]]$ as follows (in the sequel, dropping the subscript k for the sake of brevity):

$$\begin{aligned}
&\mathbb{E}[z_I\mathrm{sign}[z_R] - z_R\mathrm{sign}[z_I]\,|\,\psi]\\
&\approx \mathbb{E}\,[(x_I\cos(\psi)+x_R\sin(\psi))\mathrm{sign}[x_R\cos(\psi)-x_I\sin(\psi)]\\
&\quad -(x_R\cos(\psi)-x_I\sin(\psi))\mathrm{sign}[x_I\cos(\psi)+x_R\sin(\psi)]\,|\,\psi]\\
&\qquad \{\text{using Taylor's series approximations, below we find}\}\\
&\approx \mathbb{E}\,[(x_I+x_R\psi)\mathrm{sign}[x_R-x_I\psi]-(x_R-x_I\psi)\mathrm{sign}[x_I+x_R\psi]|\psi]\\
&\qquad \{\text{assuming } (x_R\psi \ll x_I) \text{ and } (x_I\psi \ll x_R)\}\\
&\approx \psi\,\mathbb{E}[|x_I|+|x_R|] = 2\psi\,\mathbb{E}|x_R|. \qquad (85)
\end{aligned}$$

Next we require to evaluate $\mathbb{E}[\mathcal{S}^2\,|\,\psi] = \mathbb{E}|z|^2 - 2\,\mathbb{E}[|z_R\,z_I|\,|\,\psi]$. Due to independency between signal and noise, we obtain $\mathbb{E}|z|^2 = \mathbb{E}|x|^2 + 2\sigma_e^2$. Ignoring noise and assuming again a small offset, we easily obtain $\mathbb{E}[|z_R\,z_I|\,|\,\psi] \approx (1-\psi^2)\mathbb{E}|x_R\,x_I|$.

$$\begin{aligned}
&\mathbb{E}[|z_R\,z_I|\,|\,\psi]\\
&= \mathbb{E}\big[\big|(x_R\cos(\psi)-x_I\sin(\psi)+e_R)(x_I\cos(\psi)\\
&\quad + x_R\sin(\psi)+e_I)\big|\big|\psi\big]\\
&= \mathbb{E}\big[\big|(x_Rx_I(\cos(\psi)^2-\sin(\psi)^2)+(x_R^2-x_I^2)\cos(\psi)\sin(\psi)\\
&\quad + e_Re_I + (x_R\cos(\psi)-x_I\sin(\psi))e_I\\
&\quad + (x_I\cos(\psi)+x_R\sin(\psi))e_R\big|\big|\psi\big]\\
&\qquad \{\text{assuming}\cos(\psi)\approx 1, \sin(\psi)\approx\psi, \text{and } (x_R^2-x_I^2)\psi\approx 0\}\\
&\approx \mathbb{E}\big[\big|x_Rx_I(1-\psi^2)+(x_R-x_I\psi)e_I+(x_I+x_R\psi)e_R+e_Re_I\big|\,|\,\psi\big]\\
&\qquad \{\text{assuming } (x_R\psi \ll x_I) \text{ and } (x_I\psi \ll x_R)\}\\
&\approx \mathbb{E}\big[\big|x_Rx_I\big(1-\psi^2\big)+x_Re_I+x_Ie_R+e_Re_I\big|\,|\,\psi\big]\\
&\approx (1-\psi^2)\mathbb{E}\big|x_Rx_I\big|. \qquad (86)
\end{aligned}$$

Substituting these statistics $\mathbb{E}[\mathcal{S}^2\,|\,\psi]$ and $\mathbb{E}[\mathcal{S}\,|\,\psi]$ in the variance relation (32), we obtain the MSD expression for SGD-PRA.

References

[1] C.R. Johnson, P. Schniter, T.J. Endres, J.D. Behm, D.R. Brown, R.A. Casas, Blind equalization using the constant modulus criterion: a review, Proc. IEEE 86 (1998) 1927–1950.

[2] P. Campisi, G. Panci, S. Colonnese, G. Scarano, Blind phase recovery for QAM communication systems, IEEE Trans. Signal Process. 53 (4) (2005) 1348–1358.

[3] C.N. Georghiades, Blind carrier phase acquisition for QAM constellations, IEEE Trans. Commun. 45 (11) (1997) 1477–1486.

[4] A. Leclert, P. Vandamme, Universal carrier recovery loop for QASK and PSK signal sets, IEEE Trans. Commun. COM-31 (1) (1983) 130–136.

[5] W.C. Lindsey, Synchronization Systems in Communication and Control, Prentice-Hall, Englewood Cliffs, New Jersey, 1972.
[6] W.A. Sethares, J.M. Walsh, C.R. Johnson, An adaptive view of synchronization, in: Proc. IEEE Midwest Symp. Circuit and Systems, vol. 2, 2002, pp. 521–524.
[7] U. Mengali, A.N. D'Andrea, Synchronization Techniques for Digital Receivers, Plenum Press, New York, 1997.
[8] H. Meyr, M. Moeneclaey, S.A. Fechtel, Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing, John Wiley & Sons, Inc., New York, 1998.
[9] H. Kobayashi, Simultaneous adaptive estimation and decision algorithm for carrier modulated data transmission systems, IEEE Trans. Commun. Technol. COM-19 (3) (1971) 268–280.
[10] M.K. Simon, J.G. Smith, Carrier synchronization and detection of QASK signal sets, IEEE Trans. Commun. M-22 (1974) 98.
[11] H. Mathis, Blind phase synchronization for VSB signals, IEEE Trans. Broadcast. 47 (4) (2001) 340–347.
[12] D.D. Falconer, Jointly adaptive equalization and carrier recovery in two-dimensional digital communication systems, Bell Syst. Tech. J. 55 (3) (1976) 317–334.
[13] S. Moridi, H. Sari, Analysis of four decision-feedback carrier recovery loops in the presence of intersymbol interference, IEEE Trans. Commun. COM-33 (6) (1985) 543–550.
[14] H. Sari, S. Moridi, L. Desperben, P. Vandamme, Baseband equalization and carrier recovery in digital radio systems, IEEE Trans. Commun. 35 (3) (1987) 319–327.
[15] H. Sari, S. Moridi, New phase and frequency detectors for carrier recovery in PSK and QAM systems, IEEE Trans. Commun. 36 (9) (1988) 1035–1043.
[16] A. Benveniste, M. Joindot, P. Vandamme, Analyse theorique des boucles de phase numeriques en presence de canaux dispersifs, Tech. Rep., NT/MER/TSF/1, CNET Lannion, 1979.
[17] L. Franks, Carrier and bit synchronization in data communication – a tutorial review, IEEE Trans. Commun. 28 (8) (1980) 1107–1121.
[18] A. Benveniste, M. Metivier, Adaptive Algorithms and Stochastic Approximations, Springer-Verlag, Berlin, Heidelberg, 1990.
[19] O. Macchi, E. Moreau, Self-adaptive source separation using correlated signals and cross-cumulants, in: Proc. ATHOS Workshop HOS Signal Processing, vol. 73, University of Strathclyde, Edinburgh, Scotland, 1994, pp. 1–8.
[20] O. Macchi, E. Moreau, Adaptive unsupervised separation of discrete sources, Signal Process. 73 (1999) 49–66.
[21] A. Belouchrani, W. Ren, Blind carrier phase tracking with guaranteed global convergence, IEEE Trans. Signal Process. 45 (7) (1997) 1889–1894.
[22] V. Zarzoso, A.K. Nandi, Adaptive blind source separation for virtually any source probability density function, IEEE Trans. Signal Process. 48 (2) (2000) 477–488.
[23] G.J. Foschini, Equalizing without altering or detecting data, AT&T Bell Lab. Tech. J. 64 (8) (1985) 1885–1911.
[24] J.R. Barry, J.M. Kahn, Carrier synchronization for homodyne and heterodyne detection of optical quadriphase-shift keying, J. Lightwave Technol. 10 (12) (1992) 1939–1951.
[25] K. Miyauchi, T. Nagai, M. Kato, S. Ohue, Phase jitter of carrier recovery using fourth-power multiplier for QPSK and QAM transmission, IEICE Trans. Commun. E80-B (5) (1997) 755–762.
[26] A. Touzni, R.A. Casas, T.J. Endres, C.H. Strolle, S.N. Hulyalkar, Phase recovery based on minimization of single-axis constant modulus criterion: performance analysis, in: Proc. Conf. Inform. Sci. Syst, 2001, pp. 1041–1045.
[27] A. Touzni, Carrier estimation, time synchronization, and adaptive equalization with pilot estimation for 8-VSB ATSC terrestrial receivers (part 1), in: Proc. IEEE Intl. Conf. Consumer Electronics, 2003, pp. 116–117.
[28] A. Touzni, Carrier estimation, time synchronization, and adaptive equalization with pilot estimation for 8-VSB ATSC terrestrial receivers (part 2), in: Proc. IEEE Intl. Conf. Consumer Electronics, 2003, pp. 118–119.
[29] W. Chung, W.A. Sethares, C.R. Johnson, Performance analysis of blind adaptive phase offset correction based on dispersion minimization, IEEE Trans. Signal Process. 52 (6) (2004) 1750–1759.
[30] A. Touzni, R.A. Casas, T.J. Endres, S.L. Biracree, C.H. Strolle, S.N. Hulyalkar, Carrier phase estimation based on single axis constant modulus cost criterion and Bussgang criterion, U.S. Patent 7031405-B1, 2006.
[31] W. Chung, Decision-directed carrier phase offset recovery scheme for 8-VSB signals, IEEE Trans. Consum. Electron. 53 (4) (2007) 1288–1292.
[32] J.-T. Yuan, Y.-F. Huang, Blind carrier phase acquisition and tracking for 8-VSB signals, IEEE Trans. Commun. 58 (3) (2010) 769–774.
[33] J.M. Brossier, P.O. Amblard, B. Geller, Self-adaptive PLL for general QAM constellations, in: Proc. EUSIPCO, 2002.
[34] S. Landis, B.-Z. Bobrovsky, Decision directed versus non-data aided PLLs: a comparative review, IEEE Trans. Commun. 58 (4) (2010) 1256–1261.
[35] Y. Ha, W. Chung, Non-data-aided phase noise suppression scheme for CO-OFDM systems, IEEE Photonics Technol. Lett. 25 (17) (2013) 1703–1706.
[36] R. Lopez-Valcarce, Cost minimisation interpretation of fourth power phase estimator and links to multimodulus algorithm, Electron. Lett. 40 (4) (2004) 278.
[37] S. Abrar, An adaptive method for blind carrier phase recovery in a QAM receiver, in: Proc. IEEE ICIET, 2007.
[38] J.H. Ke, K.P. Zhong, Y. Gao, J.C. Cartledge, A.S. Karar, M.A. Rezania, Linewidth-tolerant and low-complexity two-stage carrier phase estimation for dual-polarization 16-QAM coherent optical fiber communications, J. Lightwave Technol. 30 (24) (2012) 3987–3992.
[39] F. Rice, M. Rice, B. Cowley, A new algorithm for 16-QAM carrier phase estimation using QPSK partitioning, Digit. Signal Process. 12 (1) (2002) 77–86.
[40] I. Fatadin, D. Ives, S.J. Savory, Laser linewidth tolerance for 16-QAM coherent optical systems using QPSK partitioning, IEEE Photonics Technol. Lett. 22 (9) (2010) 631–633.
[41] Y. Gao, A.P.T. Lau, S. Yan, C. Lu, Low-complexity and phase noise tolerant carrier phase estimation for dual-polarization 16-QAM systems, Opt. Express 19 (22) (2011) 21717–21729.
[42] Y. Gao, A.P.T. Lau, C. Lu, J. Wu, Y. Li, X. Kun, W. Li, T.J. Lin, Low-complexity two-stage carrier phase estimation for 16-QAM systems using QPSK partitioning and maximum likelihood detection, in: Proc. Optical Fiber Communication Conf., 2011, paper OMJ6.
[43] I. Fatadin, S.J. Savory, Compensation of frequency offset for 16-QAM optical coherent systems using QPSK partitioning, IEEE Photonics Technol. Lett. 23 (17) (2011) 1246–1248.
[44] K.P. Zhong, J.H. Ke, Y. Gao, J.C. Cartledge, Linewidth-tolerant and low-complexity two-stage carrier phase estimation based on modified QPSK partitioning for dual-polarization 16-QAM systems, J. Lightwave Technol. 31 (1) (2013) 50–57.
[45] F. Zhang, J. Wu, Y. Li, K. Xu, J. Lin, Multi-stage feed-forward optical carrier phase estimation based on QPSK partitioning for 64-QAM signals, Optik 124 (2013) 2557–2560.
[46] S.M. Bilal, V.C. Fludger, G. Bosco, Dual stage CPE for 64-QAM optical systems based on a modified QPSK-partitioning algorithm, IEEE Photonics Technol. Lett. 26 (3) (2014) 267–270.
[47] I. Fatadin, D. Ives, S.J. Savory, Carrier-phase estimation for 16-QAM optical coherent systems using QPSK partitioning with barycenter approximation, J. Lightwave Technol. 32 (13) (2014) 2420–2427.
[48] I. Fatadin, D. Ives, S.J. Savory, Carrier phase recovery for 16-QAM using QPSK partitioning and sliding window averaging, IEEE Photonics Technol. Lett. 26 (9) (2014) 854–858.
[49] Y. Chen, X. Huang, A simple three-stage carrier phase estimation algorithm for 16-QAM systems based on QPSK partitioning and maximum likelihood detection, Opt. Commun. 342 (2015) 44–50.
[50] I. Horikawa, Y. Saitos, 16-QAM carrier recovery with selective gated phase-locked loop, Electron. Commun. Jpn., Part 1, Commun. 63 (7) (1980) 75–84.
[51] A.J. Viterbi, A.N. Viterbi, Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission, IEEE Trans. Inf. Theory 29 (4) (1983) 543–551.
[52] S.M. Bilal, V.C. Fludger, G. Bosco, Multi-stage CPE algorithms for phase noise mitigation in 64-QAM optical systems, J. Lightwave Technol. 32 (17) (2014) 2973–2980.
[53] X. Su, L. Xi, X. Tang, Z. Zhang, S. Bai, W. Zhang, X. Zhang, A multistage CPE scheme based on crossed constellation transformation for M-QAM, IEEE Photonics Technol. Lett. 27 (1) (2015) 77–80.
[54] A.H. Sayed, Fundamentals of Adaptive Filtering, John Wiley & Sons, Hoboken, New Jersey, 2003.
[55] R.K. Martin, W.A. Sethares, R.C. Williamson, C.R. Johnson, Exploiting sparsity in adaptive filters, IEEE Trans. Signal Process. 50 (8) (2002) 1883–1894.
[56] J. Kivinen, M.K. Warmuth, Exponentiated gradient descent versus gradient descent for linear predictors, Inf. Comput. 132 (1) (1997) 1–64.
[57] R.E. Mahony, R.C. Williamson, Prior knowledge and preferential structures in gradient descent learning algorithms, J. Mach. Learn. Res. 1 (2001) 311–355.
[58] S. Abrar, A. Zerguine, A.K. Nandi, Blind adaptive compensation of I/Q mismatch and frequency offset in low-IF receivers, in: Proc. IEEE FIT, 2014.
[59] H.-D. Han, Z. Ding, Steepest descent algorithm implementation for multichannel blind signal recovery, IET Commun. 6 (18) (Dec. 2012) 3196–3203.
[60] P. Comon, C. Jutten, Handbook of Blind Source Separation: Independent Component Analysis and Applications, Academic Press, Kidlington, Oxford, OX5 1GB, UK, 2010.
[61] S.C. Douglas, Fixed-point algorithms for the blind separation of arbitrary complex-valued non-Gaussian signal mixtures, EURASIP J. Appl. Signal Process. 2007 (1) (2007) 83.
[62] M. Novey, T. Adali, Complex ICA by negentropy maximization, IEEE Trans. Neural Netw. 19 (4) (2008) 596–609.
[63] H. Li, T. Adali, A class of complex ICA algorithms based on the kurtosis cost function, IEEE Trans. Neural Netw. 19 (3) (2008) 408–420.
[64] M. Novey, T. Adali, On extending the complex FastICA algorithm to noncircular sources, IEEE Trans. Signal Process. 56 (5) (2008) 2148–2154.
[65] J. Arenas-Garcia, A.R. Figueiras-Vidal, A.H. Sayed, Mean-square performance of a convex combination of two adaptive filters, IEEE Trans. Signal Process. 54 (3) (2006) 1078–1090.
[66] B.K. Das, M. Chakraborty, Sparse adaptive filtering by an adaptive convex combination of the LMS and the ZA-LMS algorithms, IEEE Trans. Circuits Syst. I 61 (5) (2014) 1499–1507.

Shafayat Abrar holds a B.E. degree in electrical engineering from Nadirshaw Edulji Dinshaw University of Engineering and Technology, Karachi, Pakistan (1996) and an M.S. degree in electrical engineering from King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia (2000). He received his Ph.D. degree in electrical engineering and electronics from The University of Liverpool, Liverpool, UK (2010). He has received Best Paper Award at IEEE-INCC'04, LUMS, Pakistan. He is recipient of IEEE Communications Society Heinrich Hertz Award for the Best Communications Letters for the year 2012. The areas of his research are nonlinear systems, optimization, compressed sensing, adaptive filtering, Bayesian learning, and signal processing.

Professor Azzedine Zerguine received the B.Sc. degree from Case Western Reserve University, Cleveland, OH, USA, in 1981, the M.Sc. degree from King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia, in 1990, and the Ph.D. degree from Loughborough University, Loughborough, UK, in 1996, all in electrical engineering. He is currently a Professor in the Electrical Engineering Department, KFUPM, working in the areas of signal processing and communications. He was the recipient of three Best Teaching Awards in 2000, 2005, and 2011 at KFUPM. He is presently serving as an Associate Editor of the EURASIP Journal on Advances in Signal Processing.

Professor Asoke K. Nandi received the degree of Ph.D. from the University of Cambridge (Trinity College), Cambridge (UK). He held academic positions in several universities, including Oxford (UK), Imperial College London (UK), Strathclyde (UK), and Liverpool (UK). In 2013 he moved to Brunel University (UK), to become the Chair and Head of Electronic and Computer Engineering. Professor Nandi is a Distinguished Visiting Professor at Tongji University (China), and an Adjunct Professor at University of Calgary (Canada).

His current research interests lie in the areas of signal processing and machine learning, with applications to communications, gene expression data, functional magnetic resonance data, and biomedical data. He has made many fundamental theoretical and algorithmic contributions to many aspects of signal processing and machine learning. He has much expertise in *Big Data*, dealing with heterogeneous data, and extracting information from multiple datasets obtained in different laboratories and different times. He has authored over 500 technical publications, including 200 journal papers as well as four books.

Professor Nandi is a Fellow of the Royal Academy of Engineering and also a Fellow of seven other institutions including the IEEE and the IET. Among the many awards he received are the Institute of Electrical and Electronics Engineers (USA) Heinrich Hertz Award in 2012, the Glory of Bengal Award for his outstanding achievements in scientific research in 2010, the Water Arbitration Prize of the Institution of Mechanical Engineers (UK) in 1999, and the Mountbatten Premium, Division Award of the Electronics and Communications Division, of the Institution of Electrical Engineers (UK) in 1998.